

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2006, 19:48:01 ; Search time 195 Seconds
(without alignments)
1034.230 Million cell updates/sec

Title: US-09-967-237B-2
Perfect score: 2424
Sequence: 1 MAPLCSPWLLPAPAPG.....RRGTCGVSRPAEVAETGA 459

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	100.0	459	2 AAR8058	Protein e
2	2424	100.0	459	3 AAY53228	Human MN
3	2424	100.0	459	3 AAB03005	Human MN
4	2424	100.0	459	5 AAE17175	Human RCC
5	2424	100.0	459	6 ABR58596	Human can
6	2424	100.0	459	6 ABP97744	Amino aci
7	2424	100.0	459	6 ABU56656	Lung canc
8	2424	100.0	459	6 ABU56408	Lung canc
9	2424	100.0	459	7 ADN38988	Cancer/an
10	2424	100.0	459	7 ADN39921	Cancer/an
11	2424	100.0	459	8 ADG31414	Human MN
12	2424	100.0	459	8 ADK41804	Human MN
13	2424	100.0	459	8 ADL70156	Human Car
14	2424	100.0	459	8 ADK15606	Human G25
15	2424	100.0	459	8 ADQ17644	Human sof
16	2424	100.0	459	8 ADP53999	Human car
17	2424	100.0	459	8 ADQ29712	Human col
18	2424	100.0	459	8 ADS88409	Human pro
19	2424	100.0	459	8 ADT36534	Renal cel
20	2424	100.0	459	9 ADZ64595	Human MN/
21	2419	99.8	610	4 AAB82848	Kidney ca
22	2416	99.7	459	9 AEA00140	Human TAT
23	2416	99.7	459	9 AEA00660	Human TAT
24	2007	82.8	377	3 AAY53245	MN protei

25	2007	82.8	377	3 AAB03021	Mature hu
26	2007	82.8	377	8 ADG31451	Extracell
27	2007	82.8	377	9 ADZ64600	Human MN/
28	1844.5	76.1	362	8 ADL70160	Human car
29	1600	66.0	325	8 ADL70182	Glu-tagge
30	1596	65.8	437	8 ADG31476	Murine ca
31	1576	65.0	325	8 ADL70164	Murine ca
32	1478	61.0	277	8 ADG31456	Glu-tagge
33	1398.5	57.7	429	2 AAR41746	Carbonic
34	1370	56.5	257	3 AAY53241	MN protei
35	1370	56.5	257	3 AAB03018	MN protei
36	1368	56.4	257	9 ADZ64598	Human MN
37	1364	56.3	256	2 AAR97235	Human MN
38	1340	55.3	358	8 ADG31478	Human agg
39	1210	49.9	232	8 ADG31469	N-termina
40	1108	45.7	258	8 ADG31480	Antigenic
41	763	31.5	179	8 ADL70158	Murine MN
42	562	23.2	337	3 AAY99460	Human car
43	562	23.2	337	4 AAB66209	Human PRO
44	562	23.2	337	4 AAU29157	Protein O
45	562	23.2	337	4 AAB87562	Human PRO

ALIGNMENTS

RESULT 1

AAR8058
ID AAR8058 standard; protein; 459 AA.

XX AC AAR8058;
XX DT 25-MAR-2003 (revised)
XX DT 25-JUL-1996 (first entry)
XX DE Protein encoded by MuTu putative oncogene MN.
XX KW MuTu; endogenous; cellular component; MN; HeLa cell; diagnosis;
XX KW lymphocytic choriomeningitis virus; LCMV; putative oncogene; treatment;
XX KW neoplastic; pre-neoplastic; disease; antitense therapy; antibody;
XX KW vaccine; vertebrate; immunisation; carbonic anhydrase.
XX OS Homo sapiens.

Key	Location/Qualifiers
PH Peptide	1..37
FT Peptide	/label= sig_peptide
FT Peptide	/note= "anti-MN antibody epitope"
FT Region	38..135
FT Peptide	/note= "region homologous to collagen alpha 1 chain"
FT Peptide	55..60
FT Peptide	/note= "anti-MN antibody epitope"
FT Peptide	62..67
FT Peptide	/note= "anti-MN antibody epitope"
FT Peptide	68..91
FT Peptide	/note= "anti-MN antibody epitope"
FT Peptide	127..147
FT Peptide	/note= "anti-MN antibody epitope"
FT Domain	136..391
FT Peptide	/note= "carbonic anhydrase domain"
FT Peptide	279..291
FT Peptide	/note= "anti-MN antibody epitope"
FT Region	414..433
FT Region	/note= "intracellular transmembrane region"
FT Region	434..459
FT Region	/note= "intracellular C-terminus"
FT Peptide	435..450
FT Peptide	/note= "anti-MN antibody epitope"

W09534650-A2.

21-DEC-1995.

```
XX 15-JUN-1995; 95WO-US0007628.
XX 15-JUN-1994; 94US-00260190.
PR 07-JUN-1995; 95US-00477504.
PR 07-JUN-1995; 95US-00481658.
PR 07-JUN-1995; 95US-00485049.
PR 07-JUN-1995; 95US-00485862.
PR 07-JUN-1995; 95US-00485863.
PR 07-JUN-1995; 95US-00486756.
PR 07-JUN-1995; 95US-00487077.
XX (CIBA ) CIBA CORNING DIAGNOSTICS CORP.
PA (VIRO-) INST VIROLOGY.
XX
XX Zavada J, Pastorekova S, Pastorek J;
DR WPI; 1996-049679/05.
DR N-PSDB; AAT09186.
XX
XX MN gene, protein and nucleic acid fragments - used as primers and probes
PT in the detection of MN antigens and antibodies, and in the treatment of
PT (pre)neoplastic disease.
XX
XX Claim 12; Fig 1; 102pp; English.
XX
XX The present sequence is encoded by the full length MuTu endogenous
CC cellular component, MN, cDNA clone, which was isolated from lymphocytic
CC choriomeningitis virus (LCMV) infected HeLa cells. Persistent LCMV, the
CC exogenous MuTu transmissible agent (MX), infection increases the
CC expression level of the MN gene. MN is a putative oncogene, and can
CC therefore be used in the development of prods. for the diagnosis and
CC treatment of neoplastic (NP), or pre-NP diseases. NP diseases can be
CC treated using DNA antisense to MN transcribed mRNA, anti-MN protein
CC antibodies can be used for the diagnosis NP or pre-NP diseases and a
CC vaccine contg. immunogenic amounts of the MN protein can be used to
CC immunise a vertebrate against a NP disease associated with MN antigen
CC expression. (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 459 AA;
XX
XX Query Match 100.0%; Score 2424; DB 2; Length 459;
XX Best Local Similarity 100.0%; Pred. No. 9,9e-183;
XX Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MAPLCPSPWMLPLIAPAPGLTVQLLSLLMLPVPQRLPRMQEDSPPLGGSSGGDDPL 60
DB 1 MAPLCPSPWMLPLIAPAPGLTVQLLSLLMLPVPQRLPRMQEDSPPLGGSSGGDDPL 60
QY 61 GEBDLPSSEDSPREDDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEBDLPSSEDSPREDDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPOEPQNNARHDEKGDQSHWRVGGPPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DPOEPQNNARHDEKGDQSHWRVGGPPMPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELLGFGQLPPLPELRLNNGHSVQLTTPPGLEMALGFGREYRALQLHLHWAAGRPGSEHT 240
DB 181 ELLGFGQLPPLPELRLNNGHSVQLTTPPGLEMALGFGREYRALQLHLHWAAGRPGSEHT 240
QY 241 VEGHRPPAEIHVHLSTAFARVDEALGRPGGLAVLAFLAEEGPEENSAYEQLLSRLEETA 300
DB 241 VEGHRPPAEIHVHLSTAFARVDEALGRPGGLAVLAFLAEEGPEENSAYEQLLSRLEETA 300
QY 301 EGESETQVCLDLSALLPSPFSFYQYEGSLTTPPCAQGVITVFNQTVMLSAKQHLTUS 360
DB 301 EGESETQVCLDLSALLPSPFSFYQYEGSLTTPPCAQGVITVFNQTVMLSAKQHLTUS 360
QY 361 DTLWGFQSDRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAAEFVQLNSCLAAAGDILALVF 420
DB 361 DTLWGFQSDRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAAEFVQLNSCLAAAGDILALVF 420
```

```
QY 421 GLLFAVTSVAFLVQMRQHRRTKGGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFLVQMRQHRRTKGGVSYRPAEVAETGA 459
RESULT 2
AAY53228
ID AAY53228 standard; protein; 459 AA.
XX
XX AAY53228;
XX
XX 16-JUN-2000 (first entry)
XX
XX Human MN protein SEQ ID NO:2.
XX
XX Human; MN protein; MN gene; oncogene; carbonic anhydrase; tumour;
KW oncogenesis; diagnosis; neoplastic disease; cancer; carcinoma;
KW MN/CA IX isoenzyme.
XX
XX Homo sapiens.
XX
XX US6027887-A.
XX
XX 22-FEB-2000.
XX
XX 24-JAN-1997; 97US-00787739.
XX
XX 21-OCT-1992; 92US-00964589.
PR 30-DEC-1993; 93US-00177093.
PR 15-JUN-1994; 94US-00260190.
PR 07-JUN-1995; 95US-00477504.
PR 07-JUN-1995; 95US-00481658.
PR 07-JUN-1995; 95US-00485049.
PR 07-JUN-1995; 95US-00485862.
PR 07-JUN-1995; 95US-00485863.
PR 07-JUN-1995; 95US-00486756.
PR 07-JUN-1995; 95US-00487077.
XX
XX (SLSC-) SLOVAK ACAD SCI INST VIROLOGY.
XX
XX Pastorek J, Zavada J, Pastorekova S;
XX WPI; 2000-194827/17.
XX N-PSDB; AAA16540.
XX
XX Nucleic acid based assay for diagnosing a wide variety of
PT preneoplastic/neoplastic disease comprises screening for the presence of
PT abnormal MN gene expression in a vertebrate.
XX
XX Disclosure; Fig 1; 87pp; English.
XX
XX The present invention describes a method of screening for
CC preneoplastic/neoplastic disease. The method comprises: (1) determining
CC whether abnormal MN gene expression is present in a vertebrate; and (2)
CC if abnormal MN gene expression is determined to be present in the
CC vertebrate, determining that the vertebrate has a significant risk of
CC having preneoplastic/neoplastic disease. The MN gene is an oncogene and
CC encodes an MN protein (also referred to as MN/CA IX isoenzyme). The MN
CC protein is a tumour associated carbonic anhydrase isoenzyme. The method
CC is used for detecting a wide variety of preneoplastic/neoplastic diseases
CC in a vertebrate, preferably a human. The disease detected is mammary,
CC bladder, renal, urinary tract, ovarian, uterine, cervical, endometrial,
CC vaginal, vulval, prostate, liver, lung, skin, thyroid, pancreatic,
CC testicular, brain, head and neck, mesodermal, gallbladder, rectal,
CC duodenal, jejunal, ileal, gastric, pancreatic duct, liver duct, gastric
CC mucosa, gallbladder epithelium, small intestinal mucosa, colorectal
CC mucosa, pancreatic duct epithelium or liver duct epithelium
CC preneoplastic/neoplastic disease. AAA16540 to AAA16617 and AAY53228 to
CC AAY53245 represent sequences used in the exemplification of the present
CC invention
XX
XX Sequence 459 AA;
```

Query Match 100.0%; Score 2424; DB 3; Length 459;
 Best Local Similarity 100.0%; Pred. No. 9.9e-183;
 Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWLPPLITPAPAGLTIVQLLSLLLLMPVHPQRLPRMQEDSPGLGGSGGDDPL 60
 DB 1 MAPLCSPWLPPLITPAPAGLTIVQLLSLLLLMPVHPQRLPRMQEDSPGLGGSGGDDPL 60

QY 61 GEEDLPSEEDSPREDDPPGEEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEEDLPSEEDSPREDDPPGEEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

QY 121 DPQEPQNNHRDKEGDDQSHWRVGGDPPWRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
 DB 121 DPQEPQNNHRDKEGDDQSHWRVGGDPPWRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180

QY 181 ELLGFQPLPPLPELRNRNGHSVQLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240
 DB 181 ELLGFQPLPPLPELRNRNGHSVQLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240

QY 241 VEGHRFPAAEIHVHLSAFARVDEALGRPGGLAVLAALFLEGGPEENSAYEQLLSRLEIA 300
 DB 241 VEGHRFPAAEIHVHLSAFARVDEALGRPGGLAVLAALFLEGGPEENSAYEQLLSRLEIA 300

QY 301 BEGSETQVPGLDISALLPDSFQYEGSLTTPPCAQGVITWVFNQTVMLSAKQLHTLS 360
 DB 301 BEGSETQVPGLDISALLPDSFQYEGSLTTPPCAQGVITWVFNQTVMLSAKQLHTLS 360

QY 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
 DB 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420

QY 421 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 459
 DB 421 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 459

RESULT 3

AA03005
 ID AA03005 standard; protein; 459 AA.

AC AA03005;

XX 25-SEP-2000 (first entry)

DT Human MN protein.

XX MN protein; tumour associated cell adhesion molecule; oncoprotein;
 KW proteoglycan domain; PG domain; carbonic anhydrase; CA domain;
 KW abnormal expression; neoplastic disease; cancer; gene therapy.

XX Homo sapiens.

XX WO200024913-A2.

XX 04-MAY-2000.

XX 22-OCT-1999; 99WO-US024879.

XX 23-OCT-1998; 98US-00177776.

XX 23-OCT-1998; 98US-00178115.

XX (FARB) BAYER CORP.

XX (VIRO-) INST VIROLOGY.

XX Zavada J, Pastorekova S, Pastorek J;

XX WPI; 2000-350752/30.

XX N-PSDB; AAA52459, AAA52462.

XX A molecule which specifically binds to a site on MN protein (oncoprotein)

PT and prevents adhesion of vertebrate cells to the protein, useful for

PT treating preneoplastic or neoplastic diseases such as cancer.

XX Example 1; Fig 1A-C; 154pp; English.

XX The invention relates to the inhibition of cell adhesion mediated by the
 PS oncoprotein (also known as the MN/CA IX isoenzyme or the MN/G250
 CC protein). The MN protein is a tumour-associated adhesion molecule which
 CC comprises a proteoglycan-like (PG) domain (AA03017) which contains the
 CC protein's binding site, and a carbonic anhydrase (CA) domain (AA03018).
 CC Abnormal expression of the MN protein is associated with tumorigenicity.
 CC The invention encompasses molecules (e.g., proteins and peptides) which
 CC which specifically bind to a site on the MN protein, thereby preventing
 CC adhesion of vertebrate cells to the protein in a cell adhesion assay. It
 CC also encompasses MN proteins or MN protein fragments which can be added
 CC to the extracellular environment to prevent the adhesion of vertebrate
 CC cells to each other. The invention also relates to the identification of
 CC the binding site of the MN protein and to a method of identifying a site
 CC on an MN protein to which cells adhere, comprising testing a series of
 CC overlapping peptides from the protein in a cell adhesion assay. The
 CC invention encompasses a vector comprising an expression control sequence
 CC operatively linked to a nucleic acid encoding the variable domains of a
 CC MN-specific antibody, where the domains are separated by a flexible
 CC linker peptide (AA03035) and the vector inhibits the growth of a
 CC vertebrate preneoplastic or neoplastic cell that abnormally expresses MN
 CC protein. The invention also encompasses a vector comprising a nucleic
 CC acid encoding a cytotoxic protein or peptide operatively linked to the MN
 CC gene promoter, which inhibits the growth of a vertebrate preneoplastic or
 CC neoplastic cell. Also claimed is a repressor complex that binds to the MN
 CC gene promoter (AA032473). MN proteins and peptides, MN-binding proteins
 CC and peptides, and expression vectors encoding such proteins and peptides
 CC are useful for treating patients with preneoplastic or neoplastic disease
 CC (e.g., cancers) associated with or characterised by abnormal MN
 CC expression. The present sequence represents the human MN protein

XX Sequence 459 AA;

Query Match 100.0%; Score 2424; DB 3; Length 459;

Best Local Similarity 100.0%; Pred. No. 9.9e-183;

Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWLPPLITPAPAGLTIVQLLSLLLLMPVHPQRLPRMQEDSPGLGGSGGDDPL 60

DB 1 MAPLCSPWLPPLITPAPAGLTIVQLLSLLLLMPVHPQRLPRMQEDSPGLGGSGGDDPL 60

QY 61 GEEDLPSEEDSPREDDPPGEEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

DB 61 GEEDLPSEEDSPREDDPPGEEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

QY 121 DPQEPQNNHRDKEGDDQSHWRVGGDPPWRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180

DB 121 DPQEPQNNHRDKEGDDQSHWRVGGDPPWRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180

QY 181 ELLGFQPLPPLPELRNRNGHSVQLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240

DB 181 ELLGFQPLPPLPELRNRNGHSVQLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240

QY 241 VEGHRFPAAEIHVHLSAFARVDEALGRPGGLAVLAALFLEGGPEENSAYEQLLSRLEIA 300

DB 241 VEGHRFPAAEIHVHLSAFARVDEALGRPGGLAVLAALFLEGGPEENSAYEQLLSRLEIA 300

QY 301 BEGSETQVPGLDISALLPDSFQYEGSLTTPPCAQGVITWVFNQTVMLSAKQLHTLS 360

DB 301 BEGSETQVPGLDISALLPDSFQYEGSLTTPPCAQGVITWVFNQTVMLSAKQLHTLS 360

QY 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420

DB 361 DTLWPGDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420

QY 421 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 459

DB 421 GLLFAVTSVAFVQMRQRRGTGKGVSYRPAEVAETGA 459

RESULT 4

AAE17175
ID AAE17175 standard; protein; 459 AA.
AC
XX AAE17175;
XX
DT 18-APR-2002 (first entry)
XX
DE Human RCC-associated antigen, G250 protein.
XX
XX Human; immune response; immunogenic tumour; renal cell carcinoma;
KW RCC-associated antigen; G250 protein; immunotherapy; gene therapy;
KW cancer; cytostatic; immunostimulant.
XX
OS Homo sapiens.
PN WO200198363-A2.
XX
XX 27-DEC-2001.
XX
PF 20-JUN-2001; 2001WO-NL000461.
XX
PR 20-JUN-2000; 2000US-0212669P.
XX
PA (UYN1-) UNIV NIJMEGEN.
XX
XX Visser JLM, De Vries IJM, Oosterwijk E, Figdor CG, Adema GJ;
PI WPI; 2002-139784/18.
XX
XX
XX
XX New peptides useful in preparing a composition for the treatment of
PT cancer, including renal cell carcinoma, cancer of the kidney, prostate,
PT head, neck, or gastrointestinal tract, and in eliciting immune response.
XX
XX
XX Claim 1; Page 35-36; 39pp; English.
XX
XX The patent discloses peptides which can be used to elicit an immune
CC response against a tumour, specifically against an immunogenic tumour.
CC The peptides are derived from the amino acid sequence of renal cell
CC carcinoma (RCC)-associated antigen, G250 protein. They are used in
CC immunotherapy of tumours, in particular renal cell carcinomas. Sequences
CC of the invention are useful in preparing a composition for the treatment
CC of cancer, including renal cell carcinoma, cancer of the kidney, head,
CC prostate, neck, stomach, colon, gastrointestinal tract and bladder. They
CC are also used for treating immunogenic tumours. Nucleic acid sequences
CC encoding the peptides of the invention are used in gene therapy. The
CC present sequence is human RCC-associated antigen, G250 protein
XX
XX Sequence 459 AA;
SQ

Query Match 100.0%; Score 2424; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 9.9e-183;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWLPPLIIPAPAGLTQVLLLSLLMLPVPQRLPMQEDSPILGGSSGGDDPL 60
DB 1 MAPLCSPWLPPLIIPAPAGLTQVLLLSLLMLPVPQRLPMQEDSPILGGSSGGDDPL 60
QY 61 GSEDLFSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GSEDLFSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DQEPONNAHRDKEGDQSHRWYGGDPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
DB 121 DQEPONNAHRDKEGDQSHRWYGGDPWPRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
QY 181 ELLGFOLPPLPELRLENNHSHVLTLPGLGVALGPREYRALQHLHWAAGRPGSEHT 240
DB 181 ELLGFOLPPLPELRLENNHSHVLTLPGLGVALGPREYRALQHLHWAAGRPGSEHT 240
QY 241 VEGHRRFPALIHVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLAFLAFLA 300
DB 241 VEGHRRFPALIHVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLAFLAFLA 300

QY 301 EREGSETQVPGLDISALLPSDFSRFYQEGSLTTPCAQGVITWVFNQTVMLSAKQLHTLS 360
DB 301 EREGSETQVPGLDISALLPSDFSRFYQEGSLTTPCAQGVITWVFNQTVMLSAKQLHTLS 360
QY 361 DTLWPGDSRLQNLNFRATQPLNGRVIEASFAGVDSPPRAAEPVOLNSCLAAGDILALVF 420
DB 361 DTLWPGDSRLQNLNFRATQPLNGRVIEASFAGVDSPPRAAEPVOLNSCLAAGDILALVF 420
QY 421 GLLFAVTSVAFVLVQMRQRHRRGTGKGSVSRPAEVAETGA 459
DB 421 GLLFAVTSVAFVLVQMRQRHRRGTGKGSVSRPAEVAETGA 459
RESULT 5
ID ABR58596 standard; protein; 459 AA.
XX
AC ABR58596;
XX
DT 09-JUL-2003 (first entry)
XX
DE Human cancer related protein SEQ ID NO:253.
XX
XX Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW heart disease; atherosclerosis; endometriosis.
XX
XX Homo sapiens.
XX
XX WO2003025138-A2.
XX
PD 27-MAR-2003.
XX
PF 17-SEP-2002; 2002WO-US029560.
XX
XX 17-SEP-2001; 2001US-0323469P.
PR 20-SEP-2001; 2001US-0323887P.
PR 13-NOV-2001; 2001US-0350666P.
PR 08-FEB-2002; 2002US-0355145P.
PR 08-FEB-2002; 2002US-0355257P.
PR 12-APR-2002; 2002US-0372246P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
PI Zlotnik A;
XX
XX WPI; 2003-354600/33.
DR N-PSDB; ACC72730.
XX
XX New genes that are up-regulated or down-regulated in cancers, useful as
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
PT therapeutic targets for screening drugs for treating these diseases.
XX
XX Claim 12; Page 744; 767pp; English.
XX
XX The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in

CCC	drug screening, particularly for identifying agents for treating these
CC	pathologies
XXX	
SSQ	Sequence 459 AA;
	Query Match 100.0%; Score 2424; DB 6; Length 459;
	Best Local Similarity 100.0%; Pred. No. 9.9e-183;
	Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 MAPLCPSFWPLLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQDSPLGGSSGSDPL 60
Ddb	1 MAPLCPSFWPLLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQDSPLGGSSGSDPL 60
Qy	61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPVEVKPKSEEGSKLEBDLPTVRAPG 120
Ddb	61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPVEVKPKSEEGSKLEBDLPTVRAPG 120
Qy	121 DPQEPQNNARHDKEGDDQSHWRVGGDPMPVRVSPACAGRFQSPVDIRPOLAAFCPALRPL 180
Ddb	121 DPQEPQNNARHDKEGDDQSHWRVGGDPMPVRVSPACAGRFQSPVDIRPOLAAFCPALRPL 180
Qy	181 ELLGFQLPPLPELRLRNHGSVOLTPPGLEWALGPGREYRALQHLHWGAAGRPGSEHT 240
Ddb	181 ELLGFQLPPLPELRLRNHGSVOLTPPGLEWALGPGREYRALQHLHWGAAGRPGSEHT 240
Qy	241 VEGHRFPFAETHVHVLSTAFARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLIEIA 300
Ddb	241 VEGHRFPFAETHVHVLSTAFARVDEALGRPGGLAVLAALFEEGPEENSAYEQLLSRLIEIA 300
Qy	301 EEGSETQVPGLDISALLPSDFSRYFYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
Ddb	301 EEGSETQVPGLDISALLPSDFSRYFYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
Qy	361 DTLWGPDSRLQLNFRATQPLNGRVIEASPPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
Ddb	361 DTLWGPDSRLQLNFRATQPLNGRVIEASPPAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
Qy	421 GLLFAVTSVAFVLQVNRQHRRTGKGVSYRPAEVAETGA 459
Ddb	421 GLLFAVTSVAFVLQVNRQHRRTGKGVSYRPAEVAETGA 459

RESULT 6	
ABP97744	
ID	ABP97744 standard; protein; 459 AA.
XX	
XX	
ABP97744;	
AC	
XX	
XX	
DT	28-MAY-2003 (first entry)
XX	
XX	
DE	Amino acid sequence of human CA9 polypeptide.
XX	
XX	Human; angiogenesis; wound healing; retinopathy; ischemia; inflammation;
KW	microvasculopathy; bone healing; skin inflammation; HOG3; HOG8; HOG18;
KW	follicular development; CA9; HXB; IGFBP5; HFARP; STC1; mi9-6; SSR4;
KW	cancer.

N-PSDB; ABZ77284.

Inhibiting angiogenesis for treating wound healing, retinopathy, ischemia, inflammation, microvasculopathy, bone healing, skin inflammation or follicular development by providing to a subject an antisense polynucleotide.

Claim 4; Page 47-48; 66pp; English.

The present sequence is a human CA9 polypeptide. It is used in the method of the invention. The specification describes a method modulating angiogenesis associated with wound healing, retinopathy, ischemia, inflammation, microvasculopathy, bone healing, skin inflammation or follicular development. The method comprises providing to a subject HOG3, HOG8, HOG18, CA9, HXB, IGFBP5, HFRAP, STC1, miG-6 or SR4. The methods, antisense polynucleotides, polypeptides and antibodies are useful for treating wound healing, retinopathy, ischemia, inflammation, microvasculopathy, bone healing, skin inflammation or follicular development, or cancer such as breast, colon or lung cancer, or glioblastoma

Sequence 459 AA;

RESULT 7	
ABU56656	
ID	ABU56656 standard; protein; 459 AA.
XX	
XX	
AC	ABU56656;
XX	
XX	
DT	02-APR-2003 (first entry)
XX	
DE	Lung cancer-associated polypeptide #249.
XX	
KW	Lung cancer-associated polypeptide; cytostatic; emphysema;
KW	antiinflammatory; antitastmatic; non-small cell lung cancer; atelectasis;
KW	small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

DR	N-PSDB; ABZ77284.
XX	
XX	Inhibiting angiogenesis for treating wound healing, retinopathy,
PT	ischemia, inflammation, microvasculopathy, bone healing, skin
PT	inflammation or follicular development by providing to a subject an
PT	antisense polynucleotide.
XX	
XX	Claim 4; Page 47-48; 66pp; English.
PS	
XX	
XX	The present sequence is a human CA9 polypeptide. It is used in the method
CC	of the invention. The specification describes a method modulating
CC	angiogenesis associated with wound healing, retinopathy, ischemia,
CC	inflammation, microvasculopathy, bone healing, skin inflammation or
CC	follicular development. The method comprises providing to a subject HOG3,
CC	HOG8, HOG18, CA9, HXB, IGFBP5, HFRAP, STC1, mi9-6 or SR4. The methods,
CC	antisense polynucleotides, polypeptides and antibodies are useful for
CC	treating wound healing, retinopathy, ischemia, inflammation,
CC	microvasculopathy, bone healing, skin inflammation or follicular
CC	development, or cancer such as breast, colon or lung cancer, or
CC	glioblastoma
XX	
XX	Sequence 459 AA;
SO	

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX Unidentified.
 XX WO200286443-A2.
 XX 31-OCT-2002.
 XX 18-APR-2002; 2002WO-US012476.
 XX 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX AZIZ N, Murray R;
 XX WPI; 2003-093161/08.
 XX N-PSDB; ABX76385.
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 XX expression in lung cancer.
 XX Claim 27; Page 380; 453pp; English.
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
 CC invention
 XX Sequence 459 AA;
 SQ
 Query Match 100.0%; Score 2424; DB 6; Length 459;
 Best Local Similarity 100.0%; Pred. No. 9.9e-183;
 Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAPLCPSPWLLIIPAPAGLTWQLLSLLMLMPVHPQRLPRMOEDSLGGSGSGDDPL 60
 DB 1 MAPLCPSPWLLIIPAPAGLTWQLLSLLMLMPVHPQRLPRMOEDSLGGSGSGDDPL 60
 QY 61 GEEDLPSEEDSPREEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 DB 61 GEEDLPSEEDSPREEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 QY 121 DPQEPQNNAHKDEGDDQSHRWYGGDPWPVRVSPACAGRFQSPVDTRPOLAFCPALRPL 180
 DB 121 DPQEPQNNAHKDEGDDQSHRWYGGDPWPVRVSPACAGRFQSPVDTRPOLAFCPALRPL 180
 QY 181 ELGLGFOLPPLPEURLRNNGHSVOLTLPGLLEWALGPGREYRALQLHLHWGAGRPGSGEHT 240
 DB 181 ELGLGFOLPPLPEURLRNNGHSVOLTLPGLLEWALGPGREYRALQLHLHWGAGRPGSGEHT 240

QY 241 VEGHRRPAAIHVVHLSTAFARVDEALGRPGGLAVLAALFLEEGPEENSAYEQLLSRLEEIA 300
 DB 241 VEGHRRPAAIHVVHLSTAFARVDEALGRPGGLAVLAALFLEEGPEENSAYEQLLSRLEEIA 300
 QY 301 EEGSETQVPGLDISALLPSDFSRYFOYEGSLTTPCAQGVITVFNQTVMLSAKQLHTLS 360
 DB 301 EEGSETQVPGLDISALLPSDFSRYFOYEGSLTTPCAQGVITVFNQTVMLSAKQLHTLS 360
 QY 361 DTLWGPDSRLQINFRATQPLNGRVIEASFAGVDSPPRAAPVQVQNSCLAAGDILALVF 420
 DB 361 DTLWGPDSRLQINFRATQPLNGRVIEASFAGVDSPPRAAPVQVQNSCLAAGDILALVF 420
 QY 421 GLLFAVTSVAFVQMRQRHRRGTGGVSYPRAEVAETGA 459
 DB 421 GLLFAVTSVAFVQMRQRHRRGTGGVSYPRAEVAETGA 459
 RESULT 8
 ABU56408
 ID ABU56408 standard; protein; 459 AA.
 XX
 AC ABU56408;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polypeptide #1.
 XX
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX WO200286443-A2.
 XX 31-OCT-2002.
 XX 18-APR-2002; 2002WO-US012476.
 PR 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX AZIZ N, Murray R;
 XX WPI; 2003-093161/08.
 XX N-PSDB; ABX76124.
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 XX expression in lung cancer.
 PS Claim 27; Page 189; 453pp; English.
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful

KW CA9; MN/CA9; carbonic anhydrase IX; CA IX; MN/G250; neoplastic tumour;
 KW cytosolic; cancer; tumour therapy; anti-tumour vaccine.
 XX Homo sapiens.
 XX WO2003100029-A2.
 XX PD 04-DEC-2003.
 XX XX 22-FEB-2003; 2003WO-US005136.
 XX PR 23-MAY-2002; 2002US-0383068P.
 XX PR 05-DEC-2002; 2002US-0431499P.
 XX XX (FARB) BAYER CORP.
 XX PA (VIRO-) INST VIROLOGY.
 XX PI Pastorek J, Pastorekova S, Zatovicova M, Zavada J, Ortova Gut M;
 XX DR WPI; 2004-035136/03.
 XX DR N-PSDB; ADG31413, ADG31415.
 XX XX New monoclonal antibody generated from MN/CA IX-deficient mice, where the
 XX antibody binds specifically to human tumor-associated cell adhesion
 XX protein MN/CA IX or polypeptide, useful for diagnosing, prognosing or
 XX treating cancer.
 XX XX Disclosure; SEQ ID NO 2; 156pp; English.
 XX CC This invention relates to a novel monoclonal antibody identified as the
 XX MN/CA IX specific antibody prepared in knockout mice (CA IX deficient
 XX mice). Specifically, this antibody is directed towards the MN gene, a
 XX cellular oncogene known alternatively as carbonic anhydrase 9, CA9 or
 XX MN/CA9, which encodes the MN protein that is also known as the MN/CA IX
 XX isoenzyme, carbonic anhydrase IX. CA IX or the MN/G250 protein. The
 XX present invention describes the generation of this monoclonal antibody,
 XX and immunoreactive fragments thereof, which are directed against non-
 XX immunodominant epitopes on the CA IX extracellular domain. As such, this
 XX antibody can be useful diagnostically as a marker for preneoplastic/
 XX neoplastic tumours, immunodetection methods and immunotargeting
 XX approaches. Accordingly, compositions exhibit cytostatic activity and are
 XX useful in the diagnosis, prognosis and treatment of various cancers
 XX including breast, bladder or lung cancer, in tumour therapy and in anti-
 XX tumour vaccination. This polypeptide sequence is the human MN protein of
 XX the invention.
 XX XX Sequence 459 AA;
 XX Query Match 100.0%; Score 2424; DB 8; Length 459;
 XX Best Local Similarity 100.0%; Pred. No. 9.9e-183;
 XX Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAPLCPSWLPPLIPAPAGLTQVLLSLLILMPVHPQRLPRMOEDSLPGGSGGDDPL 60
 Db 1 MAPLCPSWLPPLIPAPAGLTQVLLSLLILMPVHPQRLPRMOEDSLPGGSGGDDPL 60
 QY 61 GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 Db 61 GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
 QY 121 DPQEPQNNARHDKGDDQSHWRYGDDPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
 Db 121 DPQEPQNNARHDKGDDQSHWRYGDDPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
 QY 181 ELLGFLQPLPELRLNNGHVSQVLTLPGLMALGPGREYRALQHLHWAAGRPGSEHT 240
 Db 181 ELLGFLQPLPELRLNNGHVSQVLTLPGLMALGPGREYRALQHLHWAAGRPGSEHT 240
 QY 241 VEGHFRPAEIHVVHLSFAFARVDEALGRPGGLAVLAFAFLEEGPEENSAYEQLLSLEEIA 300
 Db 241 VEGHFRPAEIHVVHLSFAFARVDEALGRPGGLAVLAFAFLEEGPEENSAYEQLLSLEEIA 300
 QY 301 EEGSETQVPGLDISALLPSDFSRFYQYEGSLTTPPCAGQVITVFNQTVMLSAKQLHTLS.360.

Db 301 EEGSETQVPGLDISALLPSDFSRFYQYEGSLTTPPCAGQVITVFNQTVMLSAKQLHTLS 360
 QY 361 DTLWGPGDSRLQNLFRATQPLNGRVIEASFPAGVDSSPRAAEPVOLNSCLAAGDILALVF 420
 Db 361 DTLWGPGDSRLQNLFRATQPLNGRVIEASFPAGVDSSPRAAEPVOLNSCLAAGDILALVF 420
 QY 421 GLLFVAVTSVAFVQMRQRHRRGTGGVSYRPAEVAETGA 459
 Db 421 GLLFVAVTSVAFVQMRQRHRRGTGGVSYRPAEVAETGA 459
 RESULT 12
 ADK41804
 ID ADK41804 standard; protein; 459 AA.
 XX AC ADK41804;
 XX DT 06-MAY-2004 (first entry)
 XX DE Human MN protein sequence SeqID2.
 XX KW carbonic anhydrase IX; CA IX; precancerous cell; MN; cancerous cell;
 KW human; vertebrate; cytostatic; vaccine; gene therapy;
 KW renal cell carcinoma; breast cancer; colorectal cancer.
 XX OS Homo sapiens.
 XX XX WO2004005348-A1.
 XX PD 15-JAN-2004.
 XX XX 22-FEB-2003; 2003WO-US005137.
 XX PR 23-MAY-2002; 2002US-0383068P.
 XX PR 05-DEC-2002; 2002US-0431499P.
 XX XX (FARB) BAYER CORP.
 XX XX (VIRO-) INST VIROLOGY.
 XX PI Zavada J, Pastorekova S, Pastorek J, Zavadova Z;
 XX DR WPI; 2004-083500/08.
 XX DR N-PSDB; ADK41803, ADK41805.
 XX XX New soluble form of the carbonic anhydrase IX (CA IX) protein for
 XX screening, diagnosing or prognosing diseases associated with abnormal
 XX expression of CA IX protein, e.g. renal cell carcinoma, breast cancer or
 XX colorectal cancer.
 XX PT Disclosure; SEQ ID NO 2; 159pp; English.
 XX CC This invention relates to a novel soluble form of the carbonic anhydrase
 XX IX (CA IX) (or MN) protein or CA IX polypeptide which is released from
 XX precancerous and/or cancerous cells of a vertebrate into a body fluid.
 XX The invention may be useful for the development of compounds with a
 XX cytostatic activity or a vaccine whilst the disclosed sequences may be
 XX used for gene therapy. The protein and method are useful for screening,
 XX diagnosing or prognosing diseases associated with abnormal expression of
 XX carbonic anhydrase IX protein, such as precancerous and cancerous
 XX diseases like renal cell carcinoma, breast cancer or colorectal cancer.
 XX The monoclonal antibody may also be used for treating or preventing
 XX precancerous and cancerous diseases. The present sequence is that of the
 XX human MN protein of the invention.
 XX SQ Sequence 459 AA;
 Query Match 100.0%; Score 2424; DB 8; Length 459;
 Best Local Similarity 100.0%; Pred. No. 9.9e-183;
 Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAPLCPSWLPPLIPAPAGLTQVLLSLLILMPVHPQRLPRMOEDSLPGGSGGDDPL 60
 Db 1 MAPLCPSWLPPLIPAPAGLTQVLLSLLILMPVHPQRLPRMOEDSLPGGSGGDDPL 60

Db	361	DTLWGPDSRLQNFNFRATQPLNGRVIEASPPAGVDSPPRAAEVQVLSCLAAGDILALVF	420
Qy	421	GLLFVAVTSVAFVLMQRRHRTGKGVSVRPAEVAETGA	459
Db	421	GLLFVAVTSVAFVLMQRRHRTGKGVSVRPAEVAETGA	459
RESULT 14			
ID	ADK15606	standard; protein; 459 AA.	
XX	ADK15606;		
XX	03-JUN-2004	(first entry)	
XX	Human G250 tumour antigen protein.		
XX	Human; G250 tumour antigen; immunogenic peptide; CD4+; CD8+;		
KW	T-cell immune response; renal cell carcinoma; RCC; cancer; kidney;		
KW	prostate; head; neck; gastrointestinal tract; colon; stomach; bladder;		
KW	cytostatic; immunostimulant.		
XX	Homo sapiens.		
XX	US2004053391-A1.		
PN	18-MAR-2004.		
XX	11-SEP-2002; 2002US-00241814.		
XX	11-SEP-2002; 2002US-00241814.		
XX	(UYN1-) UNIV NIJMEGEN.		
XX	Visser JLM, De Vries IJM, Oosterwijk E, Figdor CG, Adema GJ;		
XX	WPI; 2004-247730/23.		
XX	New renal cell carcinoma-antigen G250-derived peptide, useful as vaccine		
XX	for treating or preventing cancer, e.g. renal cell carcinoma, cancer of		
XX	the kidney, prostate, head, neck, bladder or gastrointestinal tract, e.g.		
XX	colon or stomach.		
XX	Disclosure; SEQ ID NO 16; 17pp; English.		
XX	The present invention relates to human G250 tumour antigen derived		
XX	immunogenic peptides capable of eliciting CD4+ and CD8+ T-cell immune		
XX	responses. The immunogenic peptides can elicit an immune response against		
XX	cells that express the G250 antigen, such as those found in renal cell		
XX	carcinoma (RCC). The immunogenic peptides and/or the nucleotide sequences		
XX	that encode them are useful for the manufacture of a composition for		
XX	treating or preventing cancer, e.g. renal cell carcinoma or a cancer of		
XX	the kidney, prostate, head, neck, gastrointestinal tract or any of its		
XX	part, e.g. colon or stomach or bladder. The peptides are also useful for		
XX	the preparation of a composition for treating or preventing a tumour that		
XX	expresses the G250 antigen. The immunogenic peptides are also useful in		
XX	eliciting an immune response in a human or animal against a tumour. The		
XX	present sequence represents human G250 tumour antigen protein.		
XX	Sequence 459 AA;		
XX	Query Match	100.0%; Score 2424; DB 8; Length 459;	
XX	Best Local Similarity	100.0%; Pred. No. 9.9e-183;	
XX	Matches 459; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	MAPLCPSPWLLIPAPAGLTQVLLLSLLIMPVHPQRLPRMQDSPLGGSGGDDPL	60
Db	1	MAPLCPSPWLLIPAPAGLTQVLLLSLLIMPVHPQRLPRMQDSPLGGSGGDDPL	60
Qy	61	GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP	120
Db	61	GEEDLPSEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP	120

```
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 459 AA;
Query Match 100.0%; Score 2424; DB 8; Length 459;
Best Local Similarity 100.0%; Pred. No. 9.9e-183;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLIPAPAGLTVQLLSLLSLLLVHPVHQPRLPRMQEDSPGGSSGGDDPL 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1 MAPLCSPWMLPLIPAPAGLTVQLLSLLSLLLVHPVHQPRLPRMQEDSPGGSSGGDDPL 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 DPQEPONNAHRDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 DPQEPONNAHRDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 ELIGFQLPPLPELRLRNNGHSVQLTLPGLEMALGFGREYRALQLHLHWGAAGRPGSEHT 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 ELIGFQLPPLPELRLRNNGHSVQLTLPGLEMALGFGREYRALQLHLHWGAAGRPGSEHT 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 VEGHRRPAAIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLAFLA 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 VEGHRRPAAIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLAFLA 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 EGGSETQVPLGLDISALLPSDFSRYFOYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 EGGSETQVPLGLDISALLPSDFSRYFOYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 361 DTLWGPDSRLQLNFRATQPLNGRVIEASFPAGVDSPPRAAEPVQLNSCIAAGDILALVF 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 361 DTLWGPDSRLQLNFRATQPLNGRVIEASFPAGVDSPPRAAEPVQLNSCIAAGDILALVF 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 421 GLLFAVTSVAFLVQMRROHRRGTGGVSYRPAEVAETGA 459
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 421 GLLFAVTSVAFLVQMRROHRRGTGGVSYRPAEVAETGA 459
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: February 12, 2006, 19:51:27
Job time : 197 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2424	100.0	459	2	I38013	p54/59N - human
2	526	21.7	307	2	A29993	carbonate dehydrat
3	511	21.1	319	2	S71877	carbonate dehydrat
4	492	20.3	308	1	CRHU6	carbonate dehydrat
5	425	17.5	303	2	S12867	carbonate dehydrat
6	412	17.0	290	2	JN0576	carbonate dehydrat
7	409	16.9	264	1	CRHU7	carbonate dehydrat
8	402.5	16.6	260	2	T08463	carbonate dehydrat
9	391	16.1	259	1	CRB02	carbonate dehydrat
10	380.5	15.7	260	1	CRWS2	carbonate dehydrat
11	379.5	15.7	260	1	CRHU2	carbonate dehydrat
12	379	15.6	259	1	CRSH2	carbonate dehydrat
13	374.5	15.4	259	1	CRRB2	carbonate dehydrat
14	368.5	15.2	260	2	JH0527	carbonate dehydrat
15	366.5	15.1	260	1	CRM01R	carbonate dehydrat
16	364.5	15.0	260	2	JC2580	carbonate dehydrat
17	363	15.0	261	2	A26344	carbonate dehydrat
18	362.5	15.0	261	2	JN0836	carbonate dehydrat
19	359.5	14.8	260	2	I52551	carbonate dehydrat
20	359.5	14.8	261	1	CRHU1	carbonate dehydrat
21	359.5	14.8	261	2	JN0835	carbonate dehydrat
22	354	14.6	260	1	CRH01D	carbonate dehydrat
23	354	14.6	260	2	A43641	carbonic anhydrase
24	347	14.3	309	2	I51900	carbonate dehydrat
25	343.5	14.2	260	1	CRHU3	carbonate dehydrat
26	342	14.1	304	2	I59261	carbonate dehydrat
27	338.5	14.0	259	2	A22612	hypothetical prote
28	333	13.7	334	2	T16772	carbonate dehydrat
29	328.5	13.6	305	1	CRHU5	carbonate dehydrat

A;Title: Comparison of the 5' regions of human and mouse carbonic anhydrase II genes and
A;Reference number: A90655; MUID:8607780; PMID:3000449
A;Accession: B23202
A:Molecule type: DNA
A;Residues: 1-77 <VE2>
A;Cross-references: UNIPARC:UPI0000172F83
A;Note: The authors translated the codon CAG for residue 39 as His
R;Curtis, P.J.; Withers, E.; Demuth, D.; Watt, R.; Venta, P.J.; Tashian, R.E.
Gene 25, 325-332, 1983
A;Title: The nucleotide sequence and derived amino acid sequence of cDNA coding for mouse
carbonic dehydratase II
A;Reference number: A01143; MUID:84109569; PMID:6420240
A;Accession: A01143
A:Molecule type: mRNA
A;Residues: 2-38,'H','40'-260 <CUR>
A;Cross-references: UNIPARC:UPI0000029EES; GB:K00811; GB:K00812; GB:M1830; NID:g192333
A;Note: Initiator Met not shown
B;Curtis, P.J. Chem. 258, 4459-4463, 1983
A;Title: Cloning of mouse carbonic anhydrase mRNA and its induction in mouse erythrocyte
culture
A;Reference number: A20539; MUID:83161023; PMID:6187736
A;Accession: A20539
A:Molecule type: mRNA
A;Residues: 15-178;214-240 <CU2>
A;Cross-references: UNIPARC:UPI0000172F84; UNIPARC:UPI0000172F85
R;Venta, P.J.; Montgomerie, J.C.; Wiebauer, K.; Hewett-Emmett, D.; Tashian, R.E.
Ann. N. Y. Acad. Sci. 429, 309-323, 1984
A;Title: Organization of the mouse and human carbonic anhydrase II genes.
A;Reference number: I51949; MUID:84255152; PMID:6331255
A;Accession: I51949
A;Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A;Residues: 241-260 <RES>
A;Cross-references: UNIPARC:UPI000000045F; GB:M25944; NID:g199078; PIDN:AAA39505.1; PID:
C;Genetics:
A;Gene: Car-2
A;Introns: 12/1; 78/1; 117/3; 144/1; 169/3; 221/3
C;Superfamily: carbonate dehydratase; carbonic anhydrase homology
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc
F;5-259/Domain: carbonic anhydrase homology <CAH>
F;94,96,119/Binding site: zinc (His) #status predicted

Query Match 15.7%; Score 380.5; DB 1; Length 260;
Best Local Similarity 35.6%; Pred. No. 3.5e-17;
Matches 93; Conservative 41; Mismatches 112; Indels 15; Gaps 8;

Qy 140 HWRY---GGDPWPVRVSPACAGRQSQVDIRPQLAAFCPALRPLELGLGFQPLPELRRLR 196
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 4 HWGYSKHNGPENWKKDFPIANGDQRSPVIDTATAQHDPALQPL-LISYD--KAASKSIV 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Qy 197 NNGHSVLTLPPGLEMAL---GP-GREYRALQLHLHWGAAGRPGSEHTVEGHRPPAEIHV 252
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 61 NNGHSFNVEFPDSQNVAVKLGPSLSDSVLLQFHFWGSSDGQSGSEHTVNKKKYAAELHL 120
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Qy 253 VHLSTAFARVDALGRPGGLAVLAFLKEEGPEENSAYEQLSRLLEEIAEGSETQVPGLD 312
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 121 VHMNTKYGDGFKAVQQPDGLAVIGLVFKIGP-ASQGLQKVLEALHSIKTKGRAAFANFD 179
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Qy 313 ISALLPSDFSRYFQEGSLTPPCAQGVITVFVTQMWSAKOL---HTLSDTLWGPGDS 369
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 180 PCSLLPGNL-D-YWTYPGSLTTPLLECVTWILREPITVSSEQMSHPRTLNFEEDABE 238
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Qy 370 RLQINFRATOPLNGRVIASF 390
:||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 239 AMVDNRPAQPLNKRIKASF 259
:||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :

RESULT 11
CRHU2
carbonate dehydratase (EC 4.2.1.1) II [validated] - human
N;Alternate names: carbonic anhydrase II; hepatic carbonic anhydrase
C;Species: Homo sapiens (man)
C;Date: 07-May-1981 #sequence revision 05-May-1995 #text change 09-Jul-2004
C;Accession: A27175; Accession: A92194; A92147; I37214; I51863; I51871; A01141

R;Murakami, H.; Marelich, G.P.; Grubb, J.H.; Kyle, J.W.; Sly, W.S.
Genomics 1, 159-166, 1987
A;Title: Cloning, expression, and sequence homologies of cDNA for human carbonic anhydrase II
A;Reference number: A27175; MUID: 86085190; PMID:3121496
A;Accession: A27175
A;Molecule type: mRNA
A;Residues: 1-260 <NR>
A;Cross-references: UNIPROT:P00918; UNIPARC:UPI0000110BA2; GB:J03037; NID:gl79771; PIDN:J
R;Venta, P.J.; Montgomer, C.; Hewett-Emmett, D.; Tashian, R.E.
Biochim. Biophys. Acta 826, 195-201, 1985
A;Title: Comparison of the 5' regions of human and mouse carbonic anhydrase II genes and
A;Reference number: A90655; MUID:86077780; PMID:3000449
A;Accession: A23202
A;Molecule type: DNA
A;Residues: 1-77 <VEN>
A;Cross-references: UNIPARC:UPI000016A50; GB:X03251; GB:M18100; GB:M77181; NID:gl79778;
R;Henderson, L.E.; Henriksson, D.; Nyman, P.O.
J. Biol. Chem. 251, 5457-5463, 1976
A;Title: The primary structure of human carbonic anhydrase C.
A;Reference number: A92194; MUID:77006079; PMID:823150
A;Accession: A92194
A;Molecule type: protein
A;Residues: 2-260 <HEN>
A;Cross-references: UNIPARC:UPI0000110396
R;Lin, K.T.D.; Deutsch, H.F.
J. Biol. Chem. 249, 2329-2337, 1974
A;Title: Human carbonic anhydrases. XII. The complete primary structure of the C isozyme
A;Reference number: A92147; MUID:74143468; PMID:4207120
A;Accession: A92147
A;Molecule type: protein
A;Residues: 2-260 <LIN>
A;Cross-references: UNIPARC:UPI0000110396
R;Montgomery, J.C.; Venta, P.J.; Tashian, R.E.; Hewett-Emmett, D.
Nucleic Acids Res. 15, 4687, 1987
A;Title: Nucleotide sequence of human liver carbonic anhydrase II cDNA.
A;Reference number: I37214; MUID:87231043; PMID:3108857
A;Accession: I37214
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-260 <RE3>
A;Cross-references: UNIPARC:UPI0000110BA2; EMBL:Y00339; NID:g29586; PIDN:CAA68426.1; PID
A;Experimental source: liver
A;Note: submitted to the EMBL/GenBank/DBJ databases by David Hewett-Emmett 01-JUL-1987
R;Venta, P.J.; Welty, R.J.; Johnson, T.M.; Sly, W.S.; Tashian, R.E.
Am. J. Hum. Genet. 49, 1082-1090, 1991
A;Title: Carbonic anhydrase II deficiency syndrome in a Belgian family is caused by a po
e
A;Reference number: I51863; MUID:92026087; PMID:1928091
A;Accession: I51863
A;Status: translation not shown; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-260 <RES>
A;Cross-references: UNIPARC:UPI0000110BA2; GB:M77181; NID:gl79778; PIDN:AAA51909.1; PID:
A;Note: the complete nucleotide sequence is not shown
R;Hu, P.Y.; Ernst, A.R.; Sly, W.S.; Venta, P.J.; Skaggs, L.A.; Tashian, R.E.
Am. J. Hum. Genet. 54, 602-608, 1994
A;Title: Carbonic anhydrase II deficiency: single-base deletion in exon 7 is the predomi
A;Reference number: I51871; MUID:94175074; PMID:8128957
A;Accession: I51871
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 225-242 <RE2>
A;Cross-references: UNIPARC:UPI00000004F0; GB:S69526; NID:g545850; PIDN:AAB30170.1; PID:
A;Note: wild type shown; mutant contains frameshift after residue 226
R;Eriksson, A.E.; Jones, T.A.; Liljas, A.
submitted to the Brookhaven Protein Data Bank, February 1989
A;Reference number: A50085; PDB:1CA2
A;Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 4-259
R;Liljas, A.; Kannan, K.K.; Bergsten, P.C.; Waara, I.; Fridborg, K.; Strandberg, B.; Car
Nature New Biol. 235, 131-137, 1972
A;Title: Crystal structure of human carbonic anhydrase C.
A;Reference number: A93404; MUID:72111787; PMID:4621826
A;Contents: annotation; X-ray crystallography, 2.0 angstroms

A>Note: other residues at the active site are His-64, Asn-67, Tyr-127, Leu-197, Thr-198,
C:Genetics:
A:Gene: GDB:CA2
A:Cross-references: GDB:119739; OMIM:259730
A:Map position: 8q13-8q22.1
A:Introns: 12/1, 78/1, 117/3, 148/3, 169/3, 221/3
C:Function:
A:Description: catalyzes the reversible dissociation of carbonic acid to carbon dioxide
A>Note: this form is expressed in erythrocytes and other tissues; deficiency of this for
C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
C:Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase; metalloprotein; mono
F:2-260/Product: carbonate dehydratase II #status experimental <MAT>
F:5-259/Domain: carbonic anhydrase homology <CAH>
F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental
F:94,96,119/Binding site: zinc (His) #status experimental

Query Match 15.7%; Score 379.5; DB 1; Length 260;
Best Local Similarity 34.5%; Pred. No. 4e-17;
Matches 90; Conservative 42; Mismatches 114; Indels 15; Gaps 7;

QY 140 HWRYG---GDPPWPRVSPACAGRFQSPVDIRPQLAAFCFALRPLELLGQLPPLPELR 196
DB 4 HWGKGHNGPEHWHKDFPIADGERQSPVDITHTAKYDPSLKPLSV---SYDQATSLRL 60
QY 197 NNGHSVQLTLPGLEMAL---GP-GREYRALQLHLHGAAGRPGSEHTVEGHRFP AEIHV 252
DB 61 NNGHAFNVFDDSDQKAVLKGGLDGTYRLIQFHFWGSLDQGSSEHTVDRKKYAAELHL 120
QY 253 VHLSTAFARVDEALGRPGGLAVLAALFLEEGPEENSAYEQLLSRLBEIAEGSETOVPGLD 312
DB 121 VHWNTKYDGFKAQVQDPGLAVLGIFPKVG-SAKPGLQKVVDVLDLSIKTKGSAFTNFD 179
QY 313 ISALLPSDFSRYFOYEGSLTTPPCAQGVITVFNQTVMLSAKQL---HTLSDTLMGPGDS 369
DB 180 PRGLLPESLD-YWTYPSGLTTPPLLCVTVIVLKEPISVSSEQLKFKLNFGEGEPEE 238
QY 370 RLQLNFRATQPLNGRVIASFP 390
DB 239 LMVDNWRPAQPLKNRQIVKASF 259

RESULT 12
CRSH2
carbonate dehydratase (EC 4.2.1.1) II - sheep (tentative sequence)
N/Alternate names: carbonic anhydrase II
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Mar-1981 #sequence_revision 31-Mar-1981 #text_change 09-Jul-2004
C:Accession: A01145
R:Tanis, R.J.; Ferrell, R.E.; Tashian, R.E.
Biochim. Biophys. Acta 371, 534-548, 1974
A:Title: Amino acid sequence of sheep carbonic anhydrase C.
A:Reference number: A90598; MUID:75054988; PMID:4215456
A:Accession: A01145
A:Molecule type: protein
A:Residues: 1-259 <FAN>
A:Cross-references: UNIPROT:P00922; UNIPARC:UPI0000126DEC
R:Mallet, B.; Gulian, J.M.; Sciaky, M.; Laurent, G.; Charrel, M.
Biochim. Biophys. Acta 576, 290-304, 1979
A:Title: Forne molecules multiples de l'anhydrase carbonique erythrocytaire ovine.
A:Reference number: A90628; MUID:79145542; PMID:106895
A:Contents: annotation
A>Note: one minor and three major forms were isolated chromatographically. One of these
C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
C:Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase; zinc
F:4-258/Domain: carbonic anhydrase homology <CAH>
F:1/Modified site: acetylated amino end (Ser) #status experimental
F:93,95,118/Binding site: zinc (His) #status predicted

Query Match 15.6%; Score 379; DB 1; Length 259;
Best Local Similarity 35.5%; Pred. No. 4.3e-17;
Matches 93; Conservative 41; Mismatches 112; Indels 16; Gaps 8;

QY 140 HWRYG---GDPPWPRVSPACAGRFQSPVDIRPQLAAFCFALRPLELLGQLPPLPELR 196

DB 3 HWGKGHNGPEHWHKDFPIADGERQSPVDITKAVVDPALKPALLYEQ---AASRRMV 59
QY 197 NNGHSVQLTLPGLEMAL---GP-GREYRALQLHLHGAAGRPGSEHTVEGHRFP AEIHV 252
DB 60 NNGHSFNVFDDSDQKAVLKGGLTGTYRLVQFHFWGSSDDQGSSEHTVDRKKYAAELHL 119
QY 253 VHLSTAFARVDEALGRPGGLAVLAALFLEEGPEENSAYEQLLSRLBEIAEGSETOVPGLD 312
DB 120 VHWNTKYDGFKAQVQDPGLAVLGIFPKVG-DANPALQKLVLDVLSIKTKGSAFTNFD 178
QY 313 ISALLPSDFSRYFOYEGSLTTPPCAQGVITVFNQTVMLSAKQL---HTLSDTLMGPGDS 369
DB 179 PSSLLKRLAIN-YWTYPSGLTTPPLLESVTVWLKEPTSVSSQQLKFRSLNFNASEPEL 237
QY 370 RLQLNFRATQPLNGRVIASFP 391
DB 238 LMLANWRPAQPLKNRQVRV-FP 258

RESULT 13
CRRB2
carbonate dehydratase (EC 4.2.1.1) II - rabbit (tentative sequence)
N/Alternate names: carbonic anhydrase II
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 31-Mar-1981 #sequence_revision 31-Mar-1981 #text_change 31-Mar-2000
C:Accession: A01142
R:Ferrell, R.E.; Stroup, S.K.; Tanis, R.J.; Tashian, R.E.
Biochim. Biophys. Acta 533, 1-11, 1978
A:Title: Amino acid sequence of rabbit carbonic anhydrase II.
A:Reference number: A01142; MUID:78144871; PMID:416851
A:Accession: A01142
A:Molecule type: protein
A:Residues: 1-259 <FER>
A:Cross-references: UNIPARC:UPI0000172P81
A:Note: 203-Glu was also found
C:Superfamily: carbonate dehydratase; carbonic anhydrase homology
C:Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase; zinc
F:4-258/Domain: carbonic anhydrase homology <CAH>
F:1/Modified site: acetylated amino end (Ser) #status experimental
F:93,95,118/Binding site: zinc (His) #status predicted

Query Match 15.4%; Score 374.5; DB 1; Length 259;
Best Local Similarity 34.5%; Pred. No. 8.4e-17;
Matches 90; Conservative 42; Mismatches 114; Indels 15; Gaps 8;

QY 140 HWRYG---GDPPWPRVSPACAGRFQSPVDIRPQLAAFCFALRPLELLGQLPPLPELR 196
DB 3 HWGKGHNGPEHWHKDFPIADGERQSPVDITDAAKHDPKPLR-VSYEHP--ISRRII 59
QY 197 NNGHSVQLTLPGLEMAL---GP-GREYRALQLHLHGAAGRPGSEHTVEGHRFP AEIHV 252
DB 60 NNGHSFNVFDDSDQKAVLKGGLTGTYRLIQFHFWGSSDDQGSSEHTVDRKKYAAELHL 119
QY 253 VHLSTAFARVDEALGRPGGLAVLAALFLEEGPEENSAYEQLLSRLBEIAEGSETOVPGLD 312
DB 120 VHWNTKYDGFKAQVQDPGLAVLGIFPKIG-SATPGLQKVVDVLTSLIKTKGSAFTNFD 178
QY 313 ISALLPSDFSRYFOYEGSLTTPPCAQGVITVFNQTVMLSAKQL---HTLSDTLMGPGDS 369
DB 179 PRGLLPESLD-YWTYPSGLTTPPLLCVTVIVLKEPITVSSQQLKFRSLNFNKEAPEE 237
QY 370 RLQLNFRATQPLNGRVIASFP 390
DB 238 PMVDNWRPAQPLKNRQIVKASF 258

RESULT 14
JH0527
carbonate dehydratase (EC 4.2.1.1) II - rat
N/Alternate names: carbonic anhydrase II
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C;Accession: JH0527; S20114
R;Stolte, C.A.; McGowan, M.H.; Heim, R.A.; Varia, M.; Neubauer, J.A.
Gene 109, 265-267, 1991
A;Title: Nucleotide sequence of a cDNA encoding rat brain carbonic anhydrase II and its
A;Reference number: JH0527; MUID:92112053; PMID:1765271
A;Accession: JH0527
A;Molecule type: mRNA
A;Residues: 1-260 <STO>
A;Cross-references: UNIPROT:P27139; UNIPARC:UPI0000167A51; EMBL:X58294; NID:G55837; PIDN
A;Experimental source: brain
A;Comment: Carbonate dehydratase is a monomeric zinc metalloenzyme that catalyzes the re
C;Superfamily: carbonate dehydratase; carbonic anhydrase homology
C;Keywords: carbon-oxygen lyase; hydro-lyase; zinc
F;5-259/Domain: carbonic anhydrase homology <CAH>

Query Match 15.2%; Score 368.5; DB 2; Length 260;
Best Local Similarity 34.1%; Pred. No. 2e-16;
Matches 89; Conservative 41; Mismatches 116; Indels 15; Gaps 7;
QY 140 HWRY---GGPPWVRVSPACAGRFQSPVDIRPOLAFCPALRPLELLGFGQLPPLRLR 196
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
4 HWGYSKSNPENHKEFPPIANGDRQSPVDITGTAGHDPSPQLLIC---YDKVASKSIV 60
QY 197 NNGHSVQLTLPPLGLEWAL---GP-GREYRALQLHLHWGAAGRPGSEHTVEGHRFPPEIHV 252
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
61 NNGHSFNVEFDDSDQFAVLKEGSLGSRDIQHFHWGSSDGGQSGSEHTVKKKYAAELHL 120
QY 253 VHLSTAFARDEALRPGGLAVLAALFEEGPEENSAYEQLLSLRLIEIAEGSGETQVPGLD 312
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
121 VHWTKYDGFKAQVHPDGLAVLIGFLKIGP-ASQGLQKITEALHSIKTKGKRAAFANFD 179
QY 313 ISALLPSDFSRFYQYEGSLTTPPCAQGVIVTVFNQTVMLSAKQL---HTLSDTLWPGDS 369
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
180 PCSLLPGNLD-YMTYPGSLTTPPLECTVILKEPITVSSEQMSHPKRLNFNSEGEAE 238
QY 370 RLQINFRATQPLNGRVIEASF 390
Db : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
239 LMVDNWRPAQPLKNRKIKASF 259

RESULT 15
CRMQIR
carbonate dehydratase (EC 4.2.1.1) I - rhesus macaque (tentative sequence)
N;Alternate names: carbonic anhydrase I
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 09-Jul-2004
C;Accession: A01139
R;Henriksson, D.; Tanis, R.J.; Tashian, R.E.
Biochem. Biophys. Res. Commun. 96, 135-142, 1980
A;Title: The amino acid sequence of carbonic anhydrase I from the Rhesus macaque.
A;Reference number: A01139; MUID:81062409; PMID:6776950
A;Accession: A01139
A;Molecule type: protein
A;Residues: 1-260 <HEN>
A;Cross-references: UNIPROT:P00916; UNIPARC:UPI0000126DE2
C;Superfamily: carbonate dehydratase; carbonic anhydrase homology
C;Keywords: acetylated amino end; carbon-oxygen lyase; hydro-lyase; zinc
F;5-260/Domain: carbonic anhydrase homology <CAH>
F;1/Modified site: acetylated amino end (Ala) #status experimental
F;94,96,119/Binding site: zinc (His) #status predicted

Query Match 15.1%; Score 366.5; DB 1; Length 260;
Best Local Similarity 32.6%; Pred. No. 2.7e-16;
Matches 87; Conservative 48; Mismatches 109; Indels 23; Gaps 8;
QY 132 DKGGDQSHWRYGGDPPWVRVSPACAGRFQSPVDIRPOLAFCPALRPLELLGFGQLPPLP 191
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
9 DKGPEQ-----WSKLYPIANGNNSQSPVDIKTSEAKHDTSLKPISV---SYNPAT 55
QY 192 ELRLRNNGHSVQLTLPPLGLEWAL---GP-GREYRALQLHLHWGAAGRPGSEHTVEGHRFP 247
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
56 AKEIIVNGHSHFVNFDNDRSVLKGDPFSDSYRLFQHFHWGSSNEYGSEHTVDGVKYS 115

QY 248 AEIHVVHLSTA-FARVDEALRPGGLAVLAALFEEGPEENSAYEQLLSLRLIEIAEGSGET 306
Db : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
116 SELHIVHNSAKYSSLAEAVSRADGLAVIGVLMKVG-EANPKLQKVLDAHLAIKTKGKA 174
QY 307 QVPGLDISALLPSDFSRFYQYEGSLTTPPCAQGVIVTVFNQTVMLSAKQL---HTLSDTL 363
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
175 PFTNFDPSLTLLPSSLD-FWTYSGSLTHPPLYESVTWICKESISVSSEQLAQFRSLLSNV 233
QY 364 WPGGDSRLQINFRATQPLNGRVIEASF 390
Db : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
234 EGSNPVPIQRNNRPTQPLKGRIVRAS 260

Search completed: February 12, 2006, 19:56:28
Job time : 45 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2006, 19:48:36 ; Search time 247 Seconds
(without alignments)
1311.083 Million cell updates/sec

Title: US-09-967-237B-2
Perfect score: 2424
Sequence: 1 MAPLCPSPWLLIPAPG.....RRGTGGSYRPAEVAETGA 459

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	100.0	459	1 CAH9_HUMAN	Q16790 homo sapien
2	2420	99.8	459	2 Q5T4R1_HUMAN	Q5T4R1 homo sapien
3	1596	65.8	437	1 CAH9_MOUSE	Q8VHB5 mus musculu
4	562	23.2	337	1 CAH14_HUMAN	Q9ULX7 homo sapien
5	562	23.2	337	2 Q5TB24_HUMAN	Q5TB24 homo sapien
6	558	23.0	337	1 CAH14_MOUSE	Q9WVL6 mus musculu
7	537.5	22.2	354	1 CAH12_HUMAN	Q43570 homo sapien
8	526	21.7	307	1 CAH6_SHEEP	P08060 ovis aries
9	525	21.7	343	2 Q53YB5_HUMAN	Q53YB5 homo sapien
10	519	21.4	525	2 Q6GQ30_XENLA	Q6GQ30 xenopus lae
11	515	21.2	320	1 CAH6_CAFPA	Q865C0 canis famil
12	512	21.1	355	1 CAH12_RABIT	Q9MZ30 oryctolagus
13	512	21.1	526	2 Q4SZW1_TETNG	Q4SZW1 tetraodon n
14	511	21.1	319	1 CAH6_BOVIN	P18915 bos taurus
15	507	20.9	282	2 Q4RV11_TETNG	Q4RV11 tetraodon n
16	507	20.9	313	2 Q5FC00_HUMAN	Q5FC00 homo sapien
17	498	20.5	323	2 Q58HA5_RAT	Q58HA5 rattus norv
18	496	20.5	354	1 CAH12_MOUSE	Q8CI85 mus musculu
19	493	20.3	344	2 Q8K2J1_MOUSE	Q8K2J1 mus musculu
20	488	20.1	308	1 CAH6_HUMAN	P32800 homo sapien
21	467.5	19.3	317	1 CAH6_MOUSE	P18761 mus musculu
22	465.5	19.2	317	2 Q7TNG9_MOUSE	Q7TNG9 mus musculu
23	465.5	19.2	325	2 Q80YB7_MOUSE	Q80YB7 mus musculu
24	452	18.6	305	2 Q4RR10_TETNG	Q4RR10 tetraodon n
25	448.5	18.5	229	2 Q4SL29_TETNG	Q4SL29 tetraodon n
26	446.5	18.4	314	2 Q6DBS1_BRARE	Q6DBS1 brachydanio
27	445.5	18.4	359	2 Q4SNCO_TETNG	Q4SNCO tetraodon n
28	437.5	18.0	306	2 Q4FCT9_SQUAC	Q4FCT9 squalus aca
29	435.5	18.0	268	2 Q4RRY2_TETNG	Q4RRY2 tetraodon n
30	430.5	17.8	312	2 Q7SYW3_XENLA	Q7SYW3 xenopus lae
31	427.5	17.6	260	2 Q6PFU7_BRARE	Q6PFU7 brachydanio

RESULT 1

CAH9_HUMAN	CAH9_HUMAN	STANDARD;	PRT;	459 AA.
ID	CAH9_HUMAN			
AC	Q16790;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Carbonic anhydrase IX precursor (EC 4.2.1.1) (Carbonate dehydratase IX) (CA-IX) (CAIX) (Membrane antigen MN) (P54/58N) (Renal cell carcinoma-associated antigen G250) (RCC-associated antigen G250) (pMW1).			
DE	carcinoma-associated antigen G250, MN;			
GN	Names=CA9; Synonyms=G250, MN;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.			
RC	TISSUE=Carcinoma;			
RX	MEDLINE=94366734; PubMed=8084592;			
RA	Pastorek J., Pastorekova S., Callebaut I., Mornon J.-P., Zelnik V., Opavsky R., Zlatovicova M., Liao S., Portetelle D., Stanbridge E.J., Zavada J., Burny A., Kettmann R.;			
RA	"Cloning and characterization of MN, a human tumor-associated protein with a domain homologous to carbonic anhydrase and a putative helix-loop-helix DNA binding segment.";			
RL	Oncogene 9:2877-2888(1994).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE, AND VARIANT VAL-33.			
RC	TISSUE=Renal cell carcinoma;			
RX	MEDLINE=20175484; PubMed=10709109;			
RA	Grabmaier K., Vissers J.L.M., De Weijert M.C.A., Brakenhoff R.H., Noessner E., Oosterwijk-Wakka J.C., Van Bokhoven A., Brakenhoff R.H., Noessner E., Mulders P.A., Merks G., Figdor C.G., Adema G.J., Oosterwijk E.;			
RA	"Molecular cloning and immunogenicity of renal cell carcinoma-associated antigen G250.";			
RL	Int. J. Cancer 85:865-870(2000).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	TISSUE=Colon;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B., Brownstein M.J., Uadin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			


```

DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
SQ SEQUENCE 459 AA; 49697 MW; BA67195483F0F5CE CRC64;

Query Match 99.8%; Score 2420; DB 2; Length 459;
Best Local Similarity 99.8%; Pred. No. 2.7e-134;
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSPFWLPLLPAPAPGLTVQLLSLLLMVHPQRLPRMQEDSPPLGGSSGGEDDPL 60
DB 1 MAPLCPSPFWLPLLPAPAPGLTVQLLSLLLMVHPQRLPRMQEDSPPLGGSSGGEDDPL 60
QY 61 GEEDLPSEEDSPREDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPQFPQNNNAHRDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
DB 121 DPQFPQNNNAHRDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
QY 181 ELLGFQPLPPLPELRLNNGHVSQVLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240
DB 181 ELLGFQPLPPLPELRLNNGHVSQVLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240
QY 241 VEGHRFPFAEIHVHLSSTAFARVDEALGRPGGLAVLAFLERGPENSAEYQLLSRLBEIA 300
DB 241 VEGHRFPFAEIHVHLSSTAFARVDEALGRPGGLAVLAFLERGPENSAEYQLLSRLBEIA 300
QY 301 EEGSETQVGLDIIISALLPDSFSRYFQYEGSLTTPPCAGVITWTFVNOTVMSAKQLHTLS 360
DB 301 EEGSETQVGLDIIISALLPDSFSRYFQYEGSLTTPPCAGVITWTFVNOTVMSAKQLHTLS 360
QY 361 DTLGPGDSRLQLNFRATQPLNGRVIEASFAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
DB 361 DTLGPGDSRLQLNFRATQPLNGRVIEASFAGVDSSPRAAEPVQLNSCLAAGDILALVF 420
QY 421 GLLFAVTSVAFVLQVRRQRRGTGGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFVLQVRRQRRGTGGVSYRPAEVAETGA 459

RESULT 3
CAH9 MOUSE STANDARD; PRT; 437 AA.
AC Q8VHB5; Q8K1G1; Q8VDE4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 46, Last annotation update)
DE Carbonic anhydrase IX precursor (EC 4.2.1.1) (Carbonate dehydratase IX) (CA-IX) (CAIX) (Membrane antigen MN homolog).
GN Name=Ca9; Synonyms=Ca9;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RC STRAIN=129/Ola;
RX PubMed=14604546; DOI=10.1016/j.jim.2003.08.011;
RA Zlat'ovicova M., Tarabkova K., Svasatova E., Gibadulinova A., Mucha V.,
RA Jakubickova L., Biesova Z., Rafajova M., Ortova Gut M.O., Parkkila S.,
RA Parkkila A.-K., Waheed A., Sly W.S., Horak I., Pastorek J.,
RA Pastorekova S.;
RT "Monoclonal antibodies generated in carbonic anhydrase IX-deficient mice recognize different domains of tumour-associated hypoxia-induced carbonic anhydrase IX."
RL J. Immunol. Methods 282:117-134(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RA Ortova M.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

```

```

[3]
RN NUCLEOTIDE SEQUENCE (ISOFORM 2).
RP STRAIN=ICR; TISSUE=Small intestine;
RA Wang Y.P., Yoshikawa K., Kozaki K., Miyaishi O., Nakagawa A.,
RA Muramatsu H., Kawada Y., Uchida K., Nishikawa N., Saga S.;
RT "Alternative spliced mRNA coding for MN/CA9."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Reversible hydration of carbon dioxide (By similarity).
CC -1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -1- COFACTOR: Zinc (By similarity).
CC -1- SUBUNIT: Forms oligomers linked by disulfide bonds (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8VHB5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8VHB5-2; Sequence=VSP_007409, VSP_007410;
CC -1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC
CC EMBL; AY049077; AAL14193.1; -; Genomic_DNA.
CC EMBL; AJ245857; CAC80975.1; -; mRNA.
CC EMBL; AB086322; BAC00816.1; -; mRNA.
CC HSPG; O43570; 1JDO.
CC Ensembl; ENSMUSG0000028463; Mus musculus.
CC MGI; MGI-2447188; Car9.
CC GO; GO:0005615; C:extracellular space; TAS.
CC GO; GO:0016021; C:integral to membrane; TAS.
CC InterPro; IPR001148; Euk_Coanhd.
CC PANTHER; PTHR18952; Euk_Coanhd; 1.
CC Pfam; PF00194; Carb anhydrase; 1.
CC ProDom; PD000865; Euk_Coanhd; 1.
CC PROSITE; PS00162; Euk_CO2_ANNHYDRASE; 1.
CC Alternative splicing; Glycoprotein; Lyase; Metal-binding; Signal; Transmembrane; Zinc.
KW SIGNAL 1 31 Potential.
FT CHAIN 32 437 Carbonic anhydrase IX.
FT TRANSMEM 391 411 Potential.
FT METAL 205 205 Zinc (catalytic) (By similarity).
FT METAL 207 207 Zinc (catalytic) (By similarity).
FT METAL 230 230 Zinc (catalytic) (By similarity).
FT CARBOHYD 325 325 N-linked (GlcNAc...) (Potential).
FT VARSPPLIC 282 282 G -> V (in isoform 2).
FT VARSPPLIC 283 437 Missing (in isoform 2).
FT FTID=VSP_007410.
SQ SEQUENCE 437 AA; 47265 MW; 88F23380DCD35344 CRC64;

Query Match 65.8%; Score 1596; DB 1; Length 437;
Best Local Similarity 63.4%; Pred. No. 7e-86;
Matches 320; Conservative 36; Mismatches 79; Indels 26; Gaps 6;

QY 1 MAPLCPSPFWLPLLPAPAPGLTVQLLSLLLMVHPQRLPRMQEDSPPLGGSSGGEDDPL 60
DB 1 MASLCPSPFWLPLLPAPAPGLTVQLLSLLLMVHPQRLPRMQEDSPPLGGSSGGEDDPL 55
QY 61 GEEDLPSEEDSPREDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 56 GVDVLPSEEDAPPEADP-----PDGPPPEVNSDPRMEESLGLEDLSTPEAP- 102
QY 121 DPQFPQNNNAHRDKGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
DB 103 ---EHSQGSQSGHDEKGGHSHWSYGGTLLWPQVSPACAGRFQSPVDIRPQLAAFCPALRPL 159
QY 181 ELLGFQPLPPLPELRLNNGHVSQVLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240

```

Db 160 ELLGVELQPLPELSLNNQHTVQLTLPGLKMAFGQBYRALQLHLHWGTSDDPGSEHT 219
QY 241 VEGHREPAIHHVHLSTAFARVDEALGRPGGLAVLAFLAEQPEENSAYEQLLSLREITA 300
Db 220 VNGHREPAIHHVHLSTAFSELHAGLPGGLAVLAFLAEQPEENSAYEQLLSLREITS 279
QY 301 EGSETOVGLDISALLPSDFSRFYFOYEGSLTTPPCACQGVITVFNQTVMLSAKQLHTLS 360
Db 280 EBGSKTEIFGLDVSALLPSDLSSRYRYEGSLTTPPCQGVITVFNQTVMLSAKQLHTLS 339
QY 361 DTLWPGDGRLOLNFRATOPNGRVTASGFPAGVDSPPRAAPVOLNSCLAAGDIILALVF 420
Db 340 VSLWPGDGRLOLNFRATOPNGRVTASGFPRAEDSSP---BPVHNSCTAGDILALVF 396
QY 421 GLLFAVTSVAFVLQMRQHRH--GTKGGVGYRPAEVAETGA 459
Db 397 GLLFAVTSIAFLQLRQHRHRSRGTGKRVSYSPAEMTETGA 437

RESULT 4
CAH14 HUMAN
ID CAH14 HUMAN STANDARD; PRT; 337 AA.
AC O9ULX7; O8NCF4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Carbonic anhydrase XIV precursor (SC 4.2.1.1) (Carbonate dehydratase XIV) (CA-XIV)
GN Name=CA14; ORFNames=UNQ690/PRO1335;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99443874; PubMed=10512682; DOI=10.1006/geno.1999.5938;
RA Fujikawa-Adachi K., Nishimori I., Taguchi T., Onishi S.;
RT "Human carbonic anhydrase XIV (CA14): cDNA cloning, mRNA expression, and mapping to chromosome 1.,"
RL Genomics 61:74-81(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.P., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J., Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S., Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J., Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J., Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A., Vandlen R.L., Watanabe C., Wiedand D., Woods K., Xie M.-H., Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D., Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.,"
RL Genome Res. 13:2265-2270(2003).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RX PubMed=14702039; DOI=10.1038/ngl285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata S., Watanabe S., Hiraoka S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togliya S., Komai F., Hara R., Takeuchi K., Arita M.,

Imose N., Muesashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Komatsu T., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Takahashi K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human cDNAs.,"
RL Nat. Genet. 36:40-45(2004).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S., Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y., Phelan M., Farmer A.;
RT "Cloning of human full-length cDNAs in BD Creator(TM) system donor vector.,"
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain, Lung, and Testis;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klusner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.M., Hong L., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McSwan P.J., McKernan K.J., Wang J.M., Hong L., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Gibbs R.A., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Reversible hydration of carbon dioxide.
CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: High expression in all parts of the central nervous system and lower expression in adult liver, heart, small intestine, colon, kidney, urinary bladder and skeletal muscle.
CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC -----
CC EMBL; AB025904; BAA85002.1; -; mRNA.
CC EMBL; AK358689; AAQ89052.1; -; mRNA.
CC EMBL; AK074765; BAC1191.1; -; mRNA.
CC EMBL; BT020054; AAV38857.1; -; mRNA.
CC EMBL; BC034412; AAH34412.1; -; mRNA.
CC HSSP; O43570; 1JDO.
CC SMR; Q9ULX7; 20-278.

Ensembl; ENSG00000118298; Homo sapiens.
 HGNC; HGNC:1372; CA14.
 MIM; 604832; -.
 GO; GO:0015021; C:integral to membrane; TAS.
 GO; GO:0004089; F:carbonate dehydratase activity; TAS.
 InterPro; IPR001148; Euk Coanhd.
 PANTHER; PTHR18952; Euk Coanhd; 1.
 Pfam; PF001194; Carb anhydrase; 1.
 ProDom; PD000865; Euk Coanhd; 1.
 PROSITE; PS00162; Euk CO2 ANHYDRASE; FALSE NEG.
 Glycoprotein; Lyase; Metal-binding; Signal; Transmembrane; Zinc.
 FT SIGNAL 1 15 Potential.
 FT CHAIN 16 337 Carbonic anhydrase XIV.
 FT TOPO_DOM 16 230 Extracellular (Potential).
 FT TRANSMEM 291 311 Potential.
 FT TOPO_DOM 312 337 Cytoplasmic (Potential).
 FT METAL 109 109 Zinc (catalytic) (By similarity).
 FT METAL 111 111 Zinc (catalytic) (By similarity).
 FT METAL 135 135 Zinc (catalytic) (By similarity).
 FT CARBOHYD 213 213 N-linked (GlcNAc. .) (Potential).
 FT CONFLICT 229 229 V -> A (in Ref. 3).
 SEQUENCE 337 AA; 37668 MW; 6E101C44EA70A700 CRC64;
 Query Match 23.2%; Score 562; DB 1; Length 337;
 Best Local Similarity 38.8%; Pred. No. 3.2e-25;
 Matches 128; Conservative 47; Mismatches 129; Indels 26; Gaps 7;
 QY 136 DDQSHWRY---GGDPWPVRVSPACAGRFQSPVDIRPOLAAFCPALRPLELLGFLQPLPLPE 192
 DB 17 DGGQHWTYEGPHGQDHPWASYPECGNNAQSPIDIDTSDVTFDDPLPALQPHGYDQPGTEP 76
 QY 193 LRLRNNHGSVQLTLPGLGEMALPGREYRALQHLHWAAGRP-GSEHTVEGHRPPAEIH 251
 DB 77 LDLHNNHVTQLSLPSTLYLG-GLPRKYVAAQLHLHWKQSGSGSEHQINSEATFAELH 135
 QY 252 VVHL-STAFARVDEALGRPGGLAVLAALFEEGPEPNSAYEQLLSLLEIAEGSESTQVPG 310
 DB 136 IVHYSDSDYSLSEAAERPGQLAVLGILLIEVGETKNIAVEHILSHLHVHRKDKQTSVPP 195
 QY 311 LDISALLPSDFRYFOYEGSLTPPCACQGVITVFNQTVMLSAKOLHTLSDTLWPGGDSR 370
 DB 196 FNLRELLPKQLQGYFRYNGSLTPPCYQSVLMTVFRRSQISMEQLEKLGTLFSTEEEP 255
 QY 371 LQL---NFRATQPLNGRVIEASFPAGVDSSPRAAPVQLNSCLAAAGDILALVGLLFA-- 425
 DB 256 SKLLVQNYRALQPLNQRMVFAF-----IQAGSSYTTGEMLSLGVGLVGL 302
 QY 426 --VTSVAVFLVQMRQRHRRGTGKGVSVRPAP 453
 DB 303 CLLAVVFTARKIRKRRLENRKSVVFTSAQ 332
 RESULT 6
 CAH14 MOUSE STANDARD; PRT; 337 AA.
 ID CAH14 MOUSE STANDARD; PRT; 337 AA.
 AC QSWT6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Carbonic anhydrase XIV precursor (BC 4.2.1.1) (Carbonate dehydratase XIV) (CA-XIV)
 DE Name=Ca14; Synonym=Car14, Catm;
 GN Mus musculus (Mouse);
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RC MEDLINE=99269110; PubMed=10336468; DOI=10.1074/jbc.274.22.15701;
 RA Mori K., Ogawa Y., Ebihara K., Tamura N., Tashiro K., Kuwahara T.,
 RA Mukoyama M., Sugawara A., Ozaki S., Tanaka I., Nakao K.;
 RA "Isolation and characterization of CA XIV, a novel membrane-bound
 RT carbonic anhydrase from mouse kidney";
 RL J. Biol. Chem. 274:15701-15705(1999).
 [2]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RP TISSUE=Retina;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhac N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Faney J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Reversible hydration of carbon dioxide.
 CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
 CC -!- COFACTOR: Zinc (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Most abundant in the kidney and heart,
 CC followed by the skeletal muscle, brain, lung and liver.
 CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
 CC family.

 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

 CC EMBL: AB005450; BAA78709.1; -; mRNA.
 DR EMBL: BC046995; AAH46995.1; -; mRNA.
 DR PDB: 1R35; X-ray; A/B-18-278.
 DR PDB: 1R36; X-ray; A/B-18-278.
 DR Ensembl: ENSMUSG00000038526; Mus musculus.
 DR MGI: MGI:1344341; Carli4.
 DR GO: GO:0005615; C:extracellular space; TAS.
 DR GO: GO:0016021; C:integral to membrane; TAS.
 DR InterPro: IPR001148; Euk_Coanhd.
 DR PANTHER: PTHR18952; Euk_Coanhd; 1.
 DR Pfam: PF00194; Carb_anhydrase; 1.
 DR ProDom: PD000865; Euk_Coanhd; 1.
 DR PROSITE: PS00162; EUK_CO2 ANHYDRASE; FALSE NEG.
 KW 3D-structure; Glycoprotein; Lyase; Metal-binding; Signal;
 KW Transmembrane; Zinc.
 FT SIGNAL 1 15 Potential.
 FT CHAIN 16 337 Carbonic anhydrase XIV.
 FT TOPO_DOM 16 290 Extracellular (Potential).
 FT TRANSMEM 291 311 Potential.
 FT TOPO_DOM 312 337 Cytoplasmic (Potential).
 FT METAL 109 109 Zinc (catalytic) (By similarity).
 FT METAL 111 111 Zinc (catalytic) (By similarity).
 FT METAL 135 135 Zinc (catalytic) (By similarity).
 FT CARBOHYD 213 213 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 337 AA; 37505 MW; 32F02F4DB78AC0C9 CRC64;
 Query Match 23.0%; Score 558; DB 1; Length 337;
 Best Local Similarity 39.4%; Pred. No. 5.4e-25;
 Matches 130; Conservative 42; Mismatches 130; Indels 28; Gaps 8;
 QY 136 DDQSHWY---GGDPWPVSPACAGRFQSPDIRPQLAFCFALRPLELLGFPOLPLPE 192
 DB 17 DGGHHWTYEGPHGQDHWPTSYPCGGDAQSPINQTDVIFDPLPAVQPHGYDQLGTGP 76
 QY 193 LRLRNNGHVSQVLTPLPLEMALGPGREYRALQLHLHWGAAGR-PGSEHTVEGHRFPAETH 251
 DB 77 LDLHNNGHTVQLSLPPTLHLG-GLPRKYTAAQLHLHWGQSGLEGSEHQINSEATAAELH 135
 QY 252 VVHL-STAFARVDEALGRGGLAVLAFLFEEGPEENSAYEQLLSRLEETAECSSETQVPG 310
 DB 136 VVHYSQSYSSSEAAQKQGLAVLGILIEVGTEPNPAYDHLISRLHEIRYKDKQTSVPP 195
 QY 311 LDISALLPSDFSRYFOYEGSLTPPCCAQGVITWVFNQTVMLSAKQL-----HTLSDTLWGP 366
 DB 196 FSVRELPFPOOLEGFFRYNGLSLTPPCYQSVLWTVFNRRQISMGQLEKLTLSSTEEDP 255
 QY 367 GDSRLQLNFRATQPLNGRVIEASFAGVDSPPRAEPVQVNSCLAAAGDILALVGLLFA- 425
 DB 256 SEPLVQ-NVRVPQPLNQTIFASF-----IQAGPLTYTTGEMGLGVGLIAGC 301
 QY 426 ---VTSVAVLQVWRQHRRTKGVSYRPA 452

Db 302 LCILLAVYFIQKIRKRLGNKRSVVFTSA 331
 RESULT 7
 CAH12 HUMAN
 ID CAH12 HUMAN STANDARD; PRT; 354 AA.
 AC Q43570; Q9BWG2;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Carbonic anhydrase XII precursor [EC 4.2.1.1] (Carbonate dehydratase
 DE XII) (CA-XII) (Tumor antigen HOM-RCC-3.1.3).
 GN Name-CA12;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND CHARACTERIZATION.
 RC TISSUE=Renal cell carcinoma;
 RX MEDLINE=98301622; PubMed=9636197; DOI=10.1073/pnas.95.13.7608;
 RA Tuercio O., Sahin U., Vollmar E., Siemer S., Goettert E., Seitz G.,
 RA Parkkila A.-K., Shah G.N., Grubb J.H., Pfreundschuh M., Sly W.S.;
 RT "Human carbonic anhydrase XII: cDNA cloning, expression, and
 RT chromosomal localization of a carbonic anhydrase gene that is
 RT overexpressed in some renal cell cancers.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:7608-7613(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RC TISSUE=Lung;
 RX MEDLINE=98445416; PubMed=9770531; DOI=10.1073/pnas.95.21.12596;
 RA Ivanov S.V., Kuzmin I., Wei M.-H., Pack S., Geil L., Johnson B.E.,
 RA Stanbridge E.J., Lerman M.I.;
 RT "Down-regulation of transmembrane carbonic anhydriases in renal cell
 RT carcinoma cell lines by wild-type von Hippel-Lindau transgenes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:12596-12601(1998).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
 RC TISSUE=Eye, and Kidney;
 RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klauener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 30-292.
 RX MEDLINE=21396545; PubMed=11493685; DOI=10.1073/pnas.161301298;
 RA Whittington D.A., Waheed A., Ulmasov B., Shah G.N., Grubb J.H.,
 RA Sly W.S., Christianson D.W.;
 RT "Crystal structure of the dimeric extracellular domain of human
 RT carbonic anhydrase XII, a bitopic membrane protein overexpressed in
 RT certain cancer tumor cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9545-9550(2001).
 CC -!- FUNCTION: Reversible hydration of carbon dioxide.
 CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.

-!- COFACTOR: Zinc (By similarity).

-!- ENZYME REGULATION: Inhibited by acetazolamide.

-!- SUBCELLULAR LOCATION: Type I membrane protein.

-!- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=2;

Name=1;

IsoId=O43570-1; Sequence=Displayed;

Name=2;

IsoId=O43570-2; Sequence=VSP_000772;

Note=No experimental confirmation available;

-!- TISSUE SPECIFICITY: Highly expressed in colon, kidney, prostate, intestine and activated lymphocytes. Expressed at much higher levels in the renal cell cancers than in surrounding normal kidney tissue. Moderately expressed in pancreas, ovary and testis.

-!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AF051882; AAC39789.1; -; mRNA.

EMBL; AF037335; AAC63952.1; -; mRNA.

EMBL; BC000278; AAH00278.1; -; mRNA.

EMBL; BC011691; AAH11691.1; -; mRNA.

EMBL; BC023981; AAH23981.1; -; mRNA.

PDB; 1JCZ; X-ray; A/B=30-291.

PDB; 1JDO; X-ray; A/B=30-291.

Ensembl; ENSG00000074410; Homo sapiens.

HGNc; HGNC:1371; CA12.

H-InVD; HIX0012325; --.

MIM; 603263; --.

GO; GO:0016021; C:integral to membrane; TAS.

GO; GO:0004089; F:carbonate dehydratase activity; TAS.

GO; GO:0008270; P;zinc ion binding; TAS.

InterPro; IPRO01148; Euk Coanhd.

PANTHER; PTHR18952; Euk Coanhd; 1.

Pfam; PF00194; Carb anhydrase; 1.

ProDom; PD000865; Euk Coanhd; 1.

PROSITE; PS00162; EUK_CO2_ANHYDRASE; 1.

3D-structure; Alternative splicing; Glycoprotein; Lyase; Metal-binding; Signal; Transmembrane; Zinc.

SIGNAL 1 24 Potential.

CHAIN 25 354 Carbonic anhydrase XII.

TOPO_DOM 25 301 Extracellular (Potential).

TRANSMEM 302 322 Potential.

TOPO_DOM 323 354 Cytoplasmic (Potential).

METAL 119 119 Zinc (catalytic).

METAL 121 121 Zinc (catalytic).

METAL 145 145 Zinc (catalytic).

CARBOHYD 28 28 N-linked (GlcNAc...) (Potential).

CARBOHYD 80 80 N-linked (GlcNAc...) (Potential).

CARBOHYD 162 162 N-linked (GlcNAc...) (Potential).

DISULFID 50 230 Missing (in isoform 2).

VARSPLIC 292 302 /FTID=VSP_000772.

TURN 37 38

HELIX 40 42

HELIX 43 46

STRAND 48 51

STRAND 59 60

HELIX 62 64

STRAND 65 67

TURN 69 70

TURN 75 77

STRAND 80 80

TURN 83 84

TURN 86 91

STRAND 96 99

TURN 102 103

STRAND 105 107

[illegible]

OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC	Pecora; Bovidae; Caprinae; Ovis.
OC	NCBI_TaxID=9940;
RN	[1]
RP	PROTEIN SEQUENCE.
RC	TISSUE=Saliva.
RC	MEDLINE=36294021; PubMed=3135834;
RC	Fernley R.T., Wright R.D., Coghlan J.P.;
RA	"Complete amino acid sequence of ovine salivary carbonic anhydrase.";
RA	Biochemistry 27:2815-2820(1988).
RL	-1- FUNCTION: Reversible hydration of carbon dioxide. Its role in
CC	saliva is unknown.
CC	-1- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC	-1- COFACTOR: Zinc (By similarity).
CC	-1- SUBCELLULAR LOCATION: Secreted.
CC	-1- TISSUE SPECIFICITY: Major constituent of saliva.
CC	-1- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
CC	family.
CC	-----
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use as long as its content is in no way modified and this statement is not
CC	removed.
CC	-----
DR	PIR; A29993; A29993.
DR	HSP; O43570; IUD0.
DR	InterPro; IPR001148; Euk_Coanhd.
DR	PANTHER; PTHR18952; Euk_Coanhd; 1.
DR	Pfam; PF00194; Carb anhydrase; 1.
DR	ProDom; PD000865; Euk_Coanhd; 1.
DR	PROSITE; PS00162; EUC_CO2_ANGYDRASE; 1.
KW	Direct protein sequencing; Glycoprotein; Lyase; Metal-binding; Zinc.
FT	METAL 94 94 Zinc (catalytic) (By similarity);
FT	METAL 96 96 Zinc (catalytic) (By similarity).
FT	METAL 121 121 Zinc (catalytic) (By similarity).
FT	CARBOHYD 50 50 N-linked (GlcNAc . .)
FT	CARBOHYD 239 239 N-linked (GlcNAc . .)
FT	DISULFID 25 207 Potential.
FT	VARIANT 63 63 V -> M.
FT	VARIANT 297 297 I -> M.
FT	SEQUENCE 307 AA; 35555 MW; 338682C2D45E5D6C CRC64;
SQ	-----

Query Match	21.7%	Score 526;	DB 1;	Length 307;
Best Local Similarity	43.2%	Pred. No. 3.7e-23;		
Matches 111: Conservative	40:	Mismatches 98:	Indels	8:
			Gaps	4:

141	Qy	WRYG----	GDPPWPRVSPACAGRFQSPVDTRPQLAAFCPALRPLELLGLGFQPLPEURLR	196
6	Db	WTYSEGLMDEAHWPLEYKPCGGRQSPIDLQMKVKQYNPSLRALNLTGYGLWH-GEFFPV	64	
197	Qy	NNGHSVOLTUPPGLMALGPGREYRALQLHLHWAAGR--PGSEHTVEGHRFAETHVHVH	254	
65	Db	NNGHTVQISLPSTMSMTTSDGTQYLAKQMFHWGGASSEISGSEHTVDGMRYVIEIHVHVH	124	
255	Qy	LSTAFARVDALGRPGGLAVLAFLF--EGPEENSAYEQLLSRLLEETAERGETSTQVPGLDI	313	
125	Db	YNSKYNSYEEAQKEPDGLAVLAALVEKDYENAYYSKFISHLEDIRYAGQSTVLRGLDI	184	
314	Qy	SALLPDSFSRYFOYEGSLTTPPCAQGYIWHVFQNTVMLSAKQLHTLSDDTLWGPGDSRLQL	373	
185	Db	EDMLPGDLRYYSYLSGLTTPCCTENVHFWFVADTVKLSKTQVEKLENSLLNHQNTIQN	244	
374	Qy	NFRATQPLNGRVIEASF	390	
245	Db	DYRRTOPLNHRVVEANF	261	

RESULT 9
Q53YE5 HUMAN
ID Q53YE5 HUMAN PRELIMINARY; PRT; 343 AA.
AC Q53YE5;

	13-SEP-2005	(TrEMBLrel. 31, Created)	
DT	13-SEP-2005	(TrEMBLrel. 31, Last sequence update)	
DT	13-SEP-2005	(TrEMBLrel. 31, Last annotation update)	
DE	Carbonic anhydrase XII.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
RN	{1}		
RP	NUCLEOTIDE SEQUENCE.		
RA	Kanine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,		
RA	Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,		
RA	Phefan M., Farmer A.;		
RT	"Cloning of human full-length cDNAs in BD Creator(TM) System Donor		
RL	vector.";		
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BT006656; AAP35302.1; -; mRNA.		
SQ	SEQUENCE 343 AA; 38407 MW; 87FD325255B37D18 CRC64;		
	Query Match	21.7%; Score 525; DB 2; Length 343;	
	Best Local Similarity	36.0%; Pred.No. 4.8e-23;	
	Matches 127; Conservative	45; Mismatches 141; Indels 40; Gaps 10;	
Qy	117 EAPGDPPEQNNAHRDKEGDDQSHWRY-- --GGDPMPWRVSPACAGRFQSPVDIRPQAAF 173		
Dd	19 EQFSSP-APVNG-----SKWTYFGPDGENSWSKKYPSCGLQLQSDPLHSLDIQY 67		
Qy	174 CPALRPLELLGFQLPPLPELRNRNGHSVOLTIPLPGLEMALGPGRYRALQLHLHWCAAG 233		
Dd	68 DASLTPLFEQGYNLNSANKQFLTNNGHSVKLNIPSDMHIO-GIQSRYSATQLHLHWGNPN 126		
Qy	234 RP--GSEHTVGHFPAEIHVVHL--STAFARVDALGRPGLAVLAAPLESGPENSAYEQ 291		
Dd	127 DPHGSEHTVSQHFAELHIVHNSLDYPDASTASNKSEGVLAVLIENG-SFNPSYDK 185		
Qy	292 LLRSLEIAEEGSETQVPGILDISALLPSDFSRFYQEGSLTTPPCAOGVIWTVPNQTVML 351		
Dd	186 IFSHLOHVKYKGQEAFFPGFNIELLPARTAEYYRYRGSLTTPPCNTVLWTVFRNPVQI 245		
Qy	352 SAQOLHTLSDTLM---- --GPGDSRLQNRPRATOPLNGRVIEAASFPAGVDSSPAAAEPVQL 406		
Dd	246 SQSQLLALETALYCTHMDDSPREMNINFRQVOKFDERLVYTTSFGQIIILS----- 296		
Qy	407 NSCLAAAGDILLAVFGLLFAVTSAVFLVQMRRQRHRRGTGGGVSRPAEVAFTGA 459		
Dd	297 ---LALAGILIGICI-----VVVYSIWLFRKSKIYKGDNKGVIKYKPAFKMETEA 341		

RESULT 10

ID Q6GQ30_XENLA PRELIMINARY; PRT; 525 AA.
 AC Q6GQ30;
 DT 05-JUL-2004 (T+EMBLrel. 27, Created)
 DT 05-JUL-2004 (T+EMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (T+EMBLrel. 27, Last annotation update)
 DE MGC80389 protein.
 GN Name=MGC80389;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus;
 CX NCBI_TaxID=8355;
 RN [1]
 RC NUCLEOTIDE SEQUENCE.
 RP TISSUE=liver;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zengerb B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Dlatkchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -I- COPACTOR: Binds 2 calcium ions per subunit (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted (By similarity).
DR EMBL; BC072918; AAH72918.1; -, mRNA.
DR GO; GO:0004089; F:carbonate dehydratase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006730; P:one-carbon compound metabolism; IEA.
DR InterPro; IPR001148; Euk Coanhd.
DR InterPro; IPR001759; Pentaxin.
DR Pfam; PF00194; Carb_anhydrase; 1.
DR Pfam; PF00354; Pentaxin; 1.
DR PRINTS; PR00895; PENTAXIN.
DR ProDom; PD000865; Euk Coanhd; 1.
DR ProDom; PD002153; Pentaxin; 1.
DR SMART; SM00359; PTX; 1.
DR PROSITE; PS00162; EUK_CO2_ ANHYDRASE; 1.
DR PROSITE; PS00289; PENTAXIN; UNKNOWN_1.
KW Pentaxin.
SQ SEQUENCE 525 AA; 59644 MW; DIDC237D63735578 CRC64;

Query Match 21.4%; Score 519; DB 2; Length 525;
Best Local Similarity 40.8%; Pred. No. 1.8e-22;
Matches 107; Conservative 47; Mismatches 98; Indels 10; Gaps 4;

QY 139 SH---WRYG---GDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFLPLP 191
DB 22 SHVETVYQEGELDEANWKKYPTCAKHQSPIDIQRKVRHNPQLTLELIGYDGLTG 81

QY 192 ELRLRNGHVSQVLTLPFQLEALGPGREYRALQHLHWA--AGRPGSEHTVEGHRPPAE 249
DB 82 HFVNTNGHVSQVLDLPSTMTIKGLNSLYTAVQVHLLHWGGLSETSGSEHTIDGMRYLAE 141

QY 250 IHVVHL-STAPARVDEALGRPGGLAVLAALFEEGPEENSAVEQLLSLEETAESESTOV 308
DB 142 LHVHVSNGAYKSFDEADKPNGLAVLAFLTYTNGYNTYTSDFISKLAKIRYAGQETEM 201

QY 309 PGLDISALLPDSFRYFQYEGSLTTPPCAQGVITWTFNQTWMLSAKQLHTLSDTLWPGD 368
DB 202 HTLDVMAMLPENLVNFRYDGLSTTPPCTENVLWTFVDSFVLSKTQIDLLNTLLDWHN 261

QY 369 SRLQLNFRATQPLNGRVEASF 390
DB 262 KTLANDYRHAQPHWDRTVESAF 283

CAH6 CANPA STANDARD; PRT; 320 AA.
ID CAH6 CANPA
AC Q86SC0;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Carbonic anhydrase VI precursor (EC 4.2.1.1) (Carbonate dehydratase VI) (CA-VI).
GN Name=CA6;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC Canis.
OX NCBI_TaxID=9615;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Parotid gland;
RA Murakami M.;
RT "Canine carbonic anhydrase VI (CA6), mRNA.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: Reversible hydration of carbon dioxide. Its role in saliva is unknown (By similarity).
CC -I- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -I- COPACTOR: Zinc (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; AB080972; BAC65098.1; -, mRNA.
DR HSSP; O43570; IJDO.
DR Ensembl; ENSCAPG0000019708; Canis familiaris.
DR InterPro; IPR001148; Euk Coanhd.
DR PANTHER; PTHR18952; Euk Coanhd; 1.
DR Pfam; PF00194; Carb_anhydrase; 1.
DR ProDom; PD000865; Euk Coanhd; 1.
DR PROSITE; PS00162; EUK_CO2_ ANHYDRASE; 1.
KW Glycoprotein; Lyase; Metal-binding; Signal; Zinc.
FT SIGNAL 1 17 Potential.
FT CHAIN 18 320 Carbonic anhydrase VI.
FT METAL 111 111 Zinc (catalytic) (By similarity).
FT METAL 113 113 Zinc (catalytic) (By similarity).
FT METAL 138 138 Zinc (catalytic) (By similarity).
FT CARBOHYD 256 256 N-linked (GlcNAc...) (Potential).
FT DISULFID 42 224 Potential.
SQ SEQUENCE 320 AA; 36705 MW; 6E00A8D08AA0D23F CRC64;

Query Match 21.2%; Score 515; DB 1; Length 320;
Best Local Similarity 41.6%; Pred. No. 1.7e-22;
Matches 107; Conservative 43; Mismatches 95; Indels 12; Gaps 4;

QY 137 QDSHWRYGDDPPRVSVPACAGRFQSPVDIRPQLAAFCPALRPLELLGFLPLPLRLR 196
DB 31 DQVH-----NPREYTCGTQSPIDLRKVRKQVNPSPKALKLTGYRI-QVGEFPMI 81

QY 197 NNGHVSQVLTLPFQLEALGPGREYRALQHLHWAAGR--PGSEHTVEGHRFPAAIHVVH 254
DB 82 NNGHTVQISLPPTWRWMSDGTVEIAQMHFWGSGASSEISGSEHTIDGIRFVAEIHVH 141

QY 255 LSTAFARVDEALGRPGGLAVLAALFEEGPEENSAVEQLLSLEETAESESTQVPGLDI 313
DB 142 YNSKYSVDIAQHPDGLAVLAALVKVEDYGYNTYNSFNISHLNNIRYPGOSTVLSGLDI 201

QY 314 SALLPSDFSRYFQYEGSLTTPPCAQGVITWTFNQTWMLSAKQLHTLSDTLWPGDSRLQL 373
DB 202 EDMLPENTHYTYRGSITTPPCTENVHFWLVHVRVLSISQTKWLENLSDHQNKLHS 261

QY 374 NFRATQPLNGRVEASF 390

SO SEQUENCE 526 AA; 59111 MW; 01BF8A0288331D20 CRC64;

Query Match 21.1%; Score 512; DB 2; Length 526;
Best Local Similarity 40.2%; Pred. No. 4.6e-22;
Matches 105; Conservative 47; Mismatches 101; Indels 8; Gaps 4;

QY 137 DQSHWRYG-----CDPWPVRSPACAGRFQSPVDIRPOLAFAFCPALRPLELLGFLQPLPLPE 192
DB 22 DGIHWYTKGALDQMHWPYKYPACGGKQSPIDIQQRNVRFNFDMLQLELSGDAQOQTF 81

QY 193 LRLRNGHVSQVLTLPGLGEMALGPGREYRALQLHLHWA--AGRPGSEHTVEGHRFPAEI 250
DB 82 L-MTNGHTVQIDLPPTWITEGLPGKYTAVQVHLHGWGWDLEASGAETHIDGVRYMAEL 140

QY 251 HVVHL-STAPARVDEALGRPGGLAVLAFLAELEGPEENSAYEQLLSRLEETAEBSGTQVP 309
DB 141 HVVHNSDKYSFEIARDKPDGLAVLAFFYDDGHFFENTYYSDFIANLGIKYAGOSMYIS 200

QY 310 GLDISALLPSDFSRFYQVSGSLTPPCACQGVITVFNQVMSAKQLHTLSDTLWGFGDS 369
DB 201 SLDVRSMLPENTLHFFRYQSGSLTPPCYSILMTVFDTPITLSHNQIRKLESTLMDLNDK 260

QY 370 RLQINFRATOPNGRVTEASF 390
DB 261 TLWNVYRIAQLNDRVVESSF 281

RESULT 14

CAH6_BOVIN STANDARD; PRT; 319 AA.

AC P18915; Q95322;
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Carbonic anhydrase VI precursor (EC 4.2.1.1) (Carbonate dehydratase VI) (CA-VI) (Secreted carbonic anhydrase) (Salivary carbonic anhydrase).
DE Name=CA6;
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RN TISSUE=Submandibular gland;
RC MEDLINE=96358528; PubMed=8761494;
RA Jiang W., Wolbach J.T., Gupta D.;
RT "Sequence of bovine carbonic anhydrase VI: potential recognition sites for N-acetylgalactosaminyltransferase."
RL Biochem. J. 318:291-296 (1996).
[2]
RN PROTEIN SEQUENCE OF 15-39.
RA MEDLINE=89246331; PubMed=2497732;
RA Fernley R.T., Darling P., Aldred P., Wright R.D., Coghlan J.P.;
RT "Tissue and species distribution of the secreted carbonic anhydrase isoenzyme."
RL Biochem. J. 259:91-96 (1989).
CC -!- FUNCTION: Reversible hydration of carbon dioxide. Its role in saliva is unknown.
CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Major constituent of saliva.
CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase family.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; X96503; CAA65357.1; -, mRNA.
DR PIR; S71877; S71877.
DR HSSP; O43570; LJDO.
DR InterPro; IPR001148; Euk_COanhd.
DR PANTHER; PTHR18952; Euk_COanhd; 1.
DR Pfam; PF00194; Carb_anhydrase; 1.
DR ProDom; PD000865; Euk_COanhd; 1.
DR PROSITE; PS00162; EUK CO2 ANHYDRASE; 1.
KW Direct protein sequencing; Glycoprotein; Lyase; Metal-binding; Signal;
KW Zinc..
FT SIGNAL 1 14 Carbonic anhydrase VI.
FT CHAIN 15 319 Zinc (catalytic) (By similarity).
FT METAL 106 106 Zinc (catalytic) (By similarity).
FT METAL 108 108 Zinc (catalytic) (By similarity).
FT METAL 133 133 Zinc (catalytic) (By similarity).
FT CARBOHYD 62 62 N-linked (GlcNAc..).
FT CARBOHYD 251 251 N-linked (GlcNAc..).
FT DISULFID 37 219 Potential.
FT CONFLICT 16 16 H -> S (in Ref. 2).
SQ SEQUENCE 319 AA; 37007 MW; 179884A7A9083AED CRC64;

Query Match 21.1%; Score 511; DB 1; Length 319;
Best Local Similarity 42.0%; Pred. No. 3e-22;
Matches 108; Conservative 40; Mismatches 97; Indels 12; Gaps 4;

QY 137 DQSHWRYGDPWPVRSPACAGRFQSPVDIRPOLAFAFCPALRPLELLGFLQPLPLRLR 196
DB 26 DEKHRL-----QYPCGGTRQSPIDLKKMKVYNPSLRALNLTYGLRQ-GEPPMT 76

QY 197 NNGHVSQVLTLPGLGEMALGPGREYRALQLHLHWA--AAGRPGSEHTVEGHRFPAEIHVVH 254
DB 77 NNGHTVQISLPSMRMTTDSGSQYLAKQMHFHWGDSSEISGSEHTVDGMYRIEIHVVH 136

QY 255 LSTAPARVDEALGRPGGLAVLAFLAELEGPEENSAYEQLLSRLEETAEBSGTQVPGDI 313
DB 137 YHSKYGSYEAEQNEPDGLAVLAALVEVKYAEVNTVFNQVMSAKQLHTLSDTLWGFGDSRLQL 196

QY 314 SALLPSDFSRFYQVSGSLTPPCACQGVITVFNQVMSAKQLHTLSDTLWGFGDSRLQL 373
DB 197 QDMLPGDLRYYSYLSGLTTPSTCTENVHFWVADTVKLSKTQIEKLENSLLNHQNETIQN 256

QY 374 NFRATOPNGRVTEASF 390
DB 257 NYRSTQPLNHRVVEANF 273

RESULT 15

Q4RVRI_TETNG PRELIMINARY; PRT; 282 AA.

AC Q4RVRI;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 9 SCAF14991, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG0028211001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
[1]
RN NUCLEOTIDE SEQUENCE.
RP Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N., Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A., Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B., Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S., Anhouard V., Jubin C., Castelli V., Katinka M., Vacherie B., Biemont C., Skalli Z., Cattolico L., Poulain J., de Berardinis V., Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J., Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S., Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,

RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAAE01014991; CAG07521.1; -; Genomic_DNA.
FT NON TER 282 282
SQ SEQUENCE 282 AA; 32164 MW; B4E8C0176A094201 CRC64;

Query Match 20.9%; Score 507; DB 2; Length 282;
Best Local Similarity 39.8%; Pred. No. 4.4e-22;
Matches 104; Conservative 47; Mismatches 102; Indels 8; Gaps 4;

QY 137 DQSHWRYG---GDPPWRVSPACAGRFQSPVDIRPQLAAFCPALRPLELLGFQLPPLPE 192
DB 22 DGIHWTYKEGALDOMHWPTKYPCGCKQSPIDIQORNVRFNPDMLQLSLSGYDAQQGT 81

QY 193 LRLRNGHSVQLTLPGLEMALQPGREYRALQLHLHWA--AGRPGSEHTVEGHRFPABI 250
DB 82 L-MTNNGHTVQIDLPPTWITEGLPGKYTAVQMHLHWGGWDLASGGKHTIDGVRVMAEL 140

QY 251 HVVHL-STAFARVDEALGRPGGLAVLAFLLEGPEENSAYEQLLSLLEIAEGSETQVP 309
DB 141 HVVHNSDKYKSFIEARDKPDGLAVLAFFYDGHFENTYSDFIANLGKIKYAGQSMYIS 200

QY 310 GLDISALLPSDFSRFYEGSLTTPPCAQGVITWTFNQTVMLSAKQLHTLSDTLWGPQDS 369
DB 201 SLDVRSMLPENLHFFRYQGSLLTTPCYESILWTFDPTILSHNQIRKLESTLMDLDNK 260

QY 370 RLQLNFRATQPLNGRVIEASF 390
DB 261 TLWMDYRIAQPLNDRVVESSF 281

Search completed: February 12, 2006, 19:55:39
Job time : 249 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2006, 19:55:56 ; Search time 49 Seconds
(without alignments)
774.452 Million cell updates/sec

Title: US-09-967-237B-2

Perfect score: 2424
Sequence: 1 MAPLCSPWLLIPAPAG.....RRGTGGVSRPAEVAETGA 459

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgm2_6/ptodata/1/iaa/5 COMB.pap.*
2: /cgm2_6/ptodata/1/iaa/6 COMB.pap.*
3: /cgm2_6/ptodata/1/iaa/H COMB.pap.*
4: /cgm2_6/ptodata/1/iaa/PCTUS COMB.pap.*
5: /cgm2_6/ptodata/1/iaa/RE COMB.pap.*
6: /cgm2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2424	100.0	459	1	US-08-481-658B-2
2	2424	100.0	459	1	US-08-477-504A-2
3	2424	100.0	459	1	US-08-486-756A-2
4	2424	100.0	459	1	US-08-485-862B-2
5	2424	100.0	459	2	US-08-787-739-2
6	2424	100.0	459	2	US-08-487-077A-2
7	2424	100.0	459	2	US-08-485-863A-2
8	2424	100.0	459	2	US-08-485-049D-2
9	2424	100.0	459	2	US-09-178-115-2
10	2424	100.0	459	2	US-09-177-776-2
11	2424	100.0	459	2	US-09-772-719B-2
12	2424	100.0	459	2	US-08-260-190-6
13	2424	100.0	459	2	US-09-949-016-6042
14	2420	99.8	473	2	US-09-949-016-10203
15	2227	91.9	422	2	US-08-335-469-2
16	2227	91.9	422	2	US-08-260-190-2
17	2007	82.8	377	2	US-08-787-739-87
18	2007	82.8	377	2	US-09-178-115-87
19	2007	82.8	377	2	US-09-177-776-87
20	1398.5	57.7	429	1	US-07-964-589-2
21	1398.5	57.7	429	4	PCT-US93-02024-2
22	1370	56.5	257	2	US-08-787-739-51
23	1370	56.5	257	2	US-09-178-115-51
24	1370	56.5	257	2	US-09-177-776-51
25	1364	56.3	256	1	US-08-481-658B-51
26	1364	56.3	256	1	US-08-477-504A-51
27	1364	56.3	256	1	US-08-486-756A-51

28	1364	56.3	256	1	US-08-485-862B-51	Sequence 51, Appl
29	1364	56.3	256	2	US-08-487-077A-51	Sequence 51, Appl
30	1364	56.3	256	2	US-08-485-863A-51	Sequence 51, Appl
31	1364	56.3	256	2	US-08-485-049D-51	Sequence 51, Appl
32	1364	56.3	256	2	US-09-772-719B-51	Sequence 51, Appl
33	884	36.5	170	1	US-08-481-658B-54	Sequence 54, Appl
34	884	36.5	170	1	US-08-477-504A-54	Sequence 54, Appl
35	884	36.5	170	1	US-08-486-756A-54	Sequence 54, Appl
36	884	36.5	170	1	US-08-485-862B-54	Sequence 54, Appl
37	884	36.5	170	2	US-08-487-077A-54	Sequence 54, Appl
38	884	36.5	170	2	US-08-485-863A-54	Sequence 54, Appl
39	884	36.5	170	2	US-08-485-049D-54	Sequence 54, Appl
40	884	36.5	170	2	US-09-772-719B-54	Sequence 54, Appl
41	562	23.2	337	2	US-10-012-231A-423	Sequence 423, App
42	562	23.2	337	2	US-10-015-389A-423	Sequence 423, App
43	562	23.2	337	2	US-10-006-768A-423	Sequence 423, App
44	562	23.2	337	2	US-10-015-671A-423	Sequence 423, App
45	562	23.2	337	2	US-10-015-393A-423	Sequence 423, App

ALIGNMENTS

RESULT 1
US-08-481-658B-2
; Sequence 2, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA: US/08/481.658B
; APPLICATION NUMBER: US/08/481.658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: First 37 amino acids represent
; DESCRIPTION: signal peptide, and remaining amino acids
; DESCRIPTION: represent mature protein
US-08-481-658B-2

Query Match 100.0%; Score 2424; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;


```
/ SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION NUMBER: US/09/178,115
/ APPLICATION NUMBER: US/08/485,049D
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/260,190
/ FILING DATE: 15-JUN-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Lauder, Leona L.
/ REGISTRATION NUMBER: 30,863
/ REFERENCE/DOCKET NUMBER: D-0021.3E
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-981-2034
/ TELEFAX: 415-981-0332
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 459 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ DESCRIPTION: First 37 amino acids represent
/ DESCRIPTION: signal peptide, and remaining amino acids
/ DESCRIPTION: represent mature protein
/ US-08-485-049D-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPLCSPWLPPLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Db 1 MAPLCSPWLPPLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Qy 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGE 120
Db 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGE 120
Qy 121 DPQEPQNNARHDKEGDDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPOLAAFCPALRPL 180
Db 121 DPQEPQNNARHDKEGDDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPOLAAFCPALRPL 180
Qy 181 ELLGFQPLPPLPELRLRNNGHSVQLTLPGLLEMALGPGREYRALQLHLHWAAGRPGSEHT 240
Db 181 ELLGFQPLPPLPELRLRNNGHSVQLTLPGLLEMALGPGREYRALQLHLHWAAGRPGSEHT 240
Qy 241 VEGHRFPABTHVHVLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
Db 241 VEGHRFPABTHVHVLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
Qy 301 EEGSETQVPGGLDISALLPSDFSRYFYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
Db 301 EEGSETQVPGGLDISALLPSDFSRYFYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
Qy 361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAAEFVQNLSCLAAGDILALVF 420
Db 361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAAEFVQNLSCLAAGDILALVF 420
Qy 421 GLLFAVTSVAFVQMRQRHRRGTGGVSYRPAEVAETGA 459
Db 421 GLLFAVTSVAFVQMRQRHRRGTGGVSYRPAEVAETGA 459

RESULT 9
US-09-178-115-2
/ Sequence 2, Application US/09178115
/ Patent No. 6297041
/ GENERAL INFORMATION:
/ APPLICANT: Zavada, Jan
/ APPLICANT: Pastorekova, Silvia
/ APPLICANT: Pastorek, Jaromir
/ TITLE OF INVENTION: MN gene and Protein
```

```
/ FILE REFERENCE: D-0021.5A
/ CURRENT APPLICATION NUMBER: US/09/178,115
/ CURRENT FILING DATE: 1998-10-23
/ EARLIER APPLICATION NUMBER: 09/177,776
/ EARLIER FILING DATE: 1998-10-23
/ EARLIER APPLICATION NUMBER: 08/787,739
/ EARLIER FILING DATE: 1997-01-24
/ EARLIER APPLICATION NUMBER: 08/485,049
/ EARLIER FILING DATE: 1995-06-07
/ EARLIER APPLICATION NUMBER: 08/486,756
/ EARLIER FILING DATE: 1995-06-07
/ EARLIER APPLICATION NUMBER: 08/477,504
/ EARLIER FILING DATE: 1995-06-07
/ EARLIER APPLICATION NUMBER: 08/481,658
/ EARLIER FILING DATE: 1995-06-07
/ EARLIER APPLICATION NUMBER: 08/485,862
/ EARLIER FILING DATE: 1995-06-07
/ EARLIER APPLICATION NUMBER: 08/485,863
/ EARLIER FILING DATE: 1995-06-07
/ EARLIER APPLICATION NUMBER: 08/487,077
/ EARLIER FILING DATE: 1995-06-07
/ EARLIER APPLICATION NUMBER: 08/260,190
/ EARLIER FILING DATE: 1994-06-15
/ EARLIER APPLICATION NUMBER: 08/177,093
/ EARLIER FILING DATE: 1993-12-30
/ EARLIER APPLICATION NUMBER: 07/964,589
/ EARLIER FILING DATE: 1992-10-21
/ EARLIER APPLICATION NUMBER: PV-709-92
/ EARLIER FILING DATE: 1992-03-11
/ NUMBER OF SEQ ID NOS: 116
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 459
/ TYPE: PRT
/ ORGANISM: HUMAN
/ US-09-178-115-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPLCSPWLPPLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Db 1 MAPLCSPWLPPLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Qy 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGE 120
Db 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGE 120
Qy 121 DPQEPQNNARHDKEGDDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPOLAAFCPALRPL 180
Db 121 DPQEPQNNARHDKEGDDQSHWRYGGDPPWPRVSPACAGRFQSPVDIRPOLAAFCPALRPL 180
Qy 181 ELLGFQPLPPLPELRLRNNGHSVQLTLPGLLEMALGPGREYRALQLHLHWAAGRPGSEHT 240
Db 181 ELLGFQPLPPLPELRLRNNGHSVQLTLPGLLEMALGPGREYRALQLHLHWAAGRPGSEHT 240
Qy 241 VEGHRFPABTHVHVLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
Db 241 VEGHRFPABTHVHVLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
Qy 301 EEGSETQVPGGLDISALLPSDFSRYFYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
Db 301 EEGSETQVPGGLDISALLPSDFSRYFYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
Qy 361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAAEFVQNLSCLAAGDILALVF 420
Db 361 DTLWGPDSRLQNLNFRATQPLNGRVIEASFPAGVDSPPRAAEFVQNLSCLAAGDILALVF 420
Qy 421 GLLFAVTSVAFVQMRQRHRRGTGGVSYRPAEVAETGA 459
Db 421 GLLFAVTSVAFVQMRQRHRRGTGGVSYRPAEVAETGA 459
```

361 DTLWGPGDSRLQNLFRATQPLNGRVIEASFPAGVDSPPRAAEVPQNLSCIAAGDILALVF 420
361 DTLWGPGDSRLQNLFRATQPLNGRVIEASFPAGVDSPPRAAEVPQNLSCIAAGDILALVF 420
421 GLLFAVTSVAPLVQMRQRRHRTKGGVSRPAEVAETGA 459
421 GLLFAVTSVAPLVQMRQRRHRTKGGVSRPAEVAETGA 459

RESULT 11
US-09-772-719B-2
; Sequence 2, Application US/09772719B
; Patent No. 6770438
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; Pastorekova, Silvia
; Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 465 California Street, Suite 450
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/772,719B
; FILING DATE: 30-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3A-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: First 37 amino acids represent
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-772-719B-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPLCPSPWLPPLIPAPAGLTQVLLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Db 1 MAPLCPSPWLPPLIPAPAGLTQVLLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Qy 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Qy 121 DPQEPQNNARHDKGDDQSHWRYGDPWPVRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Db 121 DPQEPQNNARHDKGDDQSHWRYGDPWPVRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Qy 181 ELLGFQLPPLPELRLNRNGHSVQLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240

361 DTLWGPGDSRLQNLFRATQPLNGRVIEASFPAGVDSPPRAAEVPQNLSCIAAGDILALVF 420
361 DTLWGPGDSRLQNLFRATQPLNGRVIEASFPAGVDSPPRAAEVPQNLSCIAAGDILALVF 420
421 GLLFAVTSVAPLVQMRQRRHRTKGGVSRPAEVAETGA 459
421 GLLFAVTSVAPLVQMRQRRHRTKGGVSRPAEVAETGA 459

RESULT 10
US-09-177-776-2
; Sequence 2, Application US/09177776A
; Patent No. 6297051
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; Pastorekova, Silvia
; Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/177,776A
; FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 459
; TYPE: PRT
; ORGANISM: HUMAN
US-09-177-776-2

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPLCPSPWLPPLIPAPAGLTQVLLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Db 1 MAPLCPSPWLPPLIPAPAGLTQVLLLSLLLLMPVHPQRLPRMQEDSPGCGSGGDDPL 60
Qy 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 61 GEEDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Qy 121 DPQEPQNNARHDKGDDQSHWRYGDPWPVRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Db 121 DPQEPQNNARHDKGDDQSHWRYGDPWPVRVSPACAGRFQSPVDIRPOLAFCPALRPL 180
Qy 181 ELLGFQLPPLPELRLNRNGHSVQLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240
Db 181 ELLGFQLPPLPELRLNRNGHSVQLTLPGLMALPGREYRALQLHLHWGAAGRPGSEHT 240
Qy 241 VEGHRFPFPAEIHVHLSTAFARVDEALGRPGGLAVLAFLERGPENSAYEQLLSRLEBIA 300
Db 241 VEGHRFPFPAEIHVHLSTAFARVDEALGRPGGLAVLAFLERGPENSAYEQLLSRLEBIA 300
Qy 301 EGGSTQVPGLDISALLPSDFSRYSFYQYEGSLTTPPCAGQVITWVFNQVWLSAKQLHTLS 360
Db 301 EGGSTQVPGLDISALLPSDFSRYSFYQYEGSLTTPPCAGQVITWVFNQVWLSAKQLHTLS 360

181 ELIGFQLPPLPELRLRNNHSHVQLTLPGLLEALGPGREYRALQLHLHWAAGRPGSEHT 240
241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIA 300
241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIA 300
301 EGSETQVPGDLISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
301 EGSETQVPGDLISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
361 DTLWPGDSSLQINFRATQPLNGRVITEASPPAGVDSSPRAAEPVQNSCLAAGDIILALVF 420
361 DTLWPGDSSLQINFRATQPLNGRVITEASPPAGVDSSPRAAEPVQNSCLAAGDIILALVF 420
421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459
421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459

RESULT 12
US-08-260-190-6
; Sequence 6, Application US/08260190A
; Patent No. 6774117
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021-2
; CURRENT APPLICATION NUMBER: US/08/260,190A
; CURRENT FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 459
; TYPE: PRT
; ORGANISM: HUMAN
US-08-260-190-6

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSMWLPLLIAPAPGLTVQLLSLLLMPVHPQRLPRMQEDSPGSGSSGDDPL 60
DB 1 MAPLCPSMWLPLLIAPAPGLTVQLLSLLLMPVHPQRLPRMQEDSPGSGSSGDDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPQEPONNAHRDKEGDDQSHRWYGGDPPMPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
DB 121 DPQEPONNAHRDKEGDDQSHRWYGGDPPMPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
QY 181 ELIGFQLPPLPELRLRNNHSHVQLTLPGLLEALGPGREYRALQLHLHWAAGRPGSEHT 240
DB 181 ELIGFQLPPLPELRLRNNHSHVQLTLPGLLEALGPGREYRALQLHLHWAAGRPGSEHT 240
QY 241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIA 300
DB 241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIA 300
QY 301 EGSETQVPGDLISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
DB 301 EGSETQVPGDLISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
QY 361 DTLWPGDSSLQINFRATQPLNGRVITEASPPAGVDSSPRAAEPVQNSCLAAGDIILALVF 420
DB 361 DTLWPGDSSLQINFRATQPLNGRVITEASPPAGVDSSPRAAEPVQNSCLAAGDIILALVF 420
421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459
421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459

QY 361 DTLWPGDSSLQINFRATQPLNGRVITEASPPAGVDSSPRAAEPVQNSCLAAGDIILALVF 420
DB 361 DTLWPGDSSLQINFRATQPLNGRVITEASPPAGVDSSPRAAEPVQNSCLAAGDIILALVF 420
QY 421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459

RESULT 13
US-09-949-016-6042
; Sequence 6042, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6042
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6042

Query Match 100.0%; Score 2424; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCPSMWLPLLIAPAPGLTVQLLSLLLMPVHPQRLPRMQEDSPGSGSSGDDPL 60
DB 1 MAPLCPSMWLPLLIAPAPGLTVQLLSLLLMPVHPQRLPRMQEDSPGSGSSGDDPL 60
QY 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
QY 121 DPQEPONNAHRDKEGDDQSHRWYGGDPPMPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
DB 121 DPQEPONNAHRDKEGDDQSHRWYGGDPPMPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
QY 181 ELIGFQLPPLPELRLRNNHSHVQLTLPGLLEALGPGREYRALQLHLHWAAGRPGSEHT 240
DB 181 ELIGFQLPPLPELRLRNNHSHVQLTLPGLLEALGPGREYRALQLHLHWAAGRPGSEHT 240
QY 241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIA 300
DB 241 VEGHRPPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAEEGPEENSAYEQLLSRLEEIA 300
QY 301 EGSETQVPGDLISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
DB 301 EGSETQVPGDLISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQLHTLS 360
QY 361 DTLWPGDSSLQINFRATQPLNGRVITEASPPAGVDSSPRAAEPVQNSCLAAGDIILALVF 420
DB 361 DTLWPGDSSLQINFRATQPLNGRVITEASPPAGVDSSPRAAEPVQNSCLAAGDIILALVF 420
421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459
421 GLLFAVTSVAFLVQMRQRHRRGTGGVSYRPAEVAETGA 459

RESULT 14

Wed Feb 15 12:41:01 2006

US-09-949-016-10203
; Sequence 10203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10203
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10203

Query Match 99.8%; Score 2420; DB 2; Length 473;
Best Local Similarity 99.8%; Pred. No. 3e-196;
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAPLCPSPWLPLLPAPAGLTVQLLSLLLMFVHPQRLPMQEDSPPLGGSGEDDPL 60
Db 15 MAPLCPSPWLPLLPAPAGLTVQLLSLLLMFVHPQRLPMQEDSPPLGGSGEDDPL 74
Qy 61 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
Db 75 GEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 134
Qy 121 DPQFQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAFCPALRPL 180
Db 135 DPQFQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACAGRFQSPVDIRPQLAFCPALRPL 194
Qy 181 ELLEGFQLPPLPELRLRNNGHVSQVLTLPGLLEWALGPGREYRALQLHLHWGAAGRPGSHT 240
Db 195 ELLEGFQLPPLPELRLRNNGHVSQVLTLPGLLEWALGPGREYRALQLHLHWGAAGRPGSHT 254
Qy 241 VEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAALFLEGPENSAAYEQLLSRLEBIA 300
Db 255 VEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAALFLEGPENSAAYEQLLSRLEBIA 314
Qy 301 BEGSETQVPGLDISALLPSDFSRYEQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
Db 315 BEGSETQVPGLDISALLPSDFSRYEQYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 374
Qy 361 DTLWGPDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVOLNSCLAAGDILALVF 420
Db 375 DTLWGPDSRLQLNFRATQPLNGRVIEASFPAGVDSSPRAAEPVOLNSCLAAGDILALVF 434
Qy 421 GLLFAVTSVAFVQMRORHRRGTGKGVSYRPAEVAETGA 459
Db 435 GLLFAVTSVAFVQMRORHRRGTGKGVSYRPAEVAETGA 473

RESULT 15
US-08-335-469-2
; Sequence 2, Application US/08335469A
; Patent No. 6004535
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: Methods to Detect and Quantify MN Protein/Polypeptide
; FILE REFERENCE: D-0021A
; CURRENT APPLICATION NUMBER: US/08/335,469A

US-08-335-469-2
; CURRENT FILING DATE: 1994-11-07
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Human
US-08-335-469-2

Query Match 91.9%; Score 2227; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 5.3e-180;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 38 QRLPRMOEDSPPLGGSGEDDPLGEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLP 97
Db 1 QRLPRMOEDSPPLGGSGEDDPLGEEDLPSEEDSPREEDPGEEDLPGEEDLPGEEDLP 60
Qy 98 VKPKSEEGSLKLEDLPTVEAPGDPQFPQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACA 157
Db 61 VKPKSEEGSLKLEDLPTVEAPGDPQFPQNNNAHRDKEGDDQSHWRYGDDPPWPRVSPACA 120
Qy 158 GRFQSPVDIRPQLAFCPALRPLLELGFQPLPPLPELRLRNNGHVSQVLTLPGLLEWALGPG 217
Db 121 GRFQSPVDIRPQLAFCPALRPLLELGFQPLPPLPELRLRNNGHVSQVLTLPGLLEWALGPG 180
Qy 218 REYRALQLHLHWGAAGRPGSEHTVEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAA 277
Db 181 REYRALQLHLHWGAAGRPGSEHTVEGHRFPAAEIHVHVLSTAFARVDEALGRPGGLAVLAA 240
Qy 278 FLEEGPENSAAYEQLLSRLEBIAEGSETQVPGLDISALLPSDFSRYEQYEGSLTTPPCA 337
Db 241 FLEEGPENSAAYEQLLSRLEBIAEGSETQVPGLDISALLPSDFSRYEQYEGSLTTPPCA 300
Qy 338 QGVITVFNQTVMLSAKQLHTLSDTLWGPDSRLQLNFRATQPLNGRVIEASFPAGVDSS 397
Db 301 QGVITVFNQTVMLSAKQLHTLSDTLWGPDSRLQLNFRATQPLNGRVIEASFPAGVDSS 360
Qy 398 PRAAEPVOLNSCLAAGDILALVFGLLFAVTSVAFVQMRORHRRGTGKGVSYRPAEVAET 457
Db 361 PRAAEPVOLNSCLAAGDILALVFGLLFAVTSVAFVQMRORHRRGTGKGVSYRPAEVAET 420
Qy 458 GA 459
Db 421 GA 422

Search completed: February 12, 2006, 19:57:24
Job time : 50 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2006, 20:07:51 ; Search time 177 Seconds
(without alignments)
1083.524 Million cell updates/sec

Title: US-09-967-237B-2

Perfect score: 2424

Sequence: 1 MAPLCSPWLLIPAPAG.....RRGTKGVSYRPAEVAETGA 459

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main.*

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2424	100.0	459	3	US-09-772-719-2
2	2424	100.0	459	3	Sequence 2, Appli
3	2424	100.0	459	4	Sequence 2, Appli
4	2424	100.0	459	4	Sequence 12, Appl
5	2424	100.0	459	4	Sequence 10, Appl
6	2424	100.0	459	4	Sequence 306, App
7	2424	100.0	459	4	Sequence 1239, Ap
8	2424	100.0	459	4	Sequence 16, Appli
9	2424	100.0	459	4	Sequence 2, Appli
10	2424	100.0	459	4	Sequence 138, App
11	2424	100.0	459	5	Sequence 461, App
12	2424	100.0	459	5	Sequence 6, Appli
13	2424	100.0	459	5	Sequence 2, Appli
14	2424	100.0	459	5	Sequence 19127, A
15	2424	100.0	459	5	Sequence 294, App
16	2424	100.0	459	5	Sequence 4846, Ap
17	2419	99.8	610	3	Sequence 1, Appli
18	2416	99.7	459	5	Sequence 92, Appl
19	2416	99.7	459	5	Sequence 92, Appl
20	2227	82.8	377	3	Sequence 2, Appli
21	2007	82.8	377	3	Sequence 87, Appl
22	2007	82.8	377	4	Sequence 7, Appli
23	2007	82.8	377	5	Sequence 10, Appl
24	1596	65.8	437	5	Sequence 90, Appl
25	1596	65.8	437	5	Sequence 18837, A
26	1596	65.8	437	5	Sequence 18838, A
27	1587	65.5	437	5	Sequence 18838, A

28	1539.5	63.5	418	5	US-10-732-923-18868
29	1478	61.0	277	5	US-10-921-590-101
30	1370	56.5	257	3	US-09-967-237-51
31	1370	56.5	257	4	US-10-723-795-5
32	1370	56.5	257	5	US-10-921-590-9
33	1364	56.3	256	3	US-09-772-719-51
34	1364	56.3	256	5	US-10-888-694-51
35	1340	55.3	358	5	US-10-921-590-92
36	1210	49.9	232	5	US-10-921-590-69
37	1210	49.9	232	5	US-10-921-590-105
38	1108	45.7	258	5	US-10-921-590-94
39	884	36.5	170	3	US-09-772-719-54
40	884	36.5	170	5	US-10-888-694-54
41	562	23.2	337	3	US-09-946-374-423
42	562	23.2	337	3	US-09-983-000A-27
43	562	23.2	337	4	US-10-006-867-74
44	562	23.2	337	4	US-10-052-586-268
45	562	23.2	337	4	US-10-063-547-74

ALIGNMENTS

RESULT 1

US-09-772-719-2
; Sequence 2, Application US/09772719
; Patent No. US20020137910A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA: US/09/772,719
; APPLICATION NUMBER: US/09/772,719
; FILING DATE: 30-JAN-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/485,049
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 459 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: First 37 amino acids represent
; DESCRIPTION: signal peptide, and remaining amino acids
; DESCRIPTION: represent mature protein
US-09-772-719-2

Db	241	VEGHRPFAETHVHLS	TAFARVDEALGRPGCLAVLAAPLEBEGPEENSAYEQLLSRLEBIA	300
Qy	301	EGSETQVPGLDISALPLSPDSR	YFQYEGSLTTPPCAQGVITWTFNQTVMLSAKQLHTLS	360
Db	301	EGSETQVPGLDISALPLSPDSR	YFQYEGSLTTPPCAQGVITWTFNQTVMLSAKQLHTLS	360
Qy	361	DTLWGPDSRLQNLNFRATOP	PLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDITLALVF	420
Db	361	DTLWGPDSRLQNLNFRATOP	PLNGRVIEASFPAGVDSSPRAAEPVQLNSCLAAGDITLALVF	420
Qy	421	GLLFAVTSVAFVQMRQRHRRGT	KGVSVYRPAEVAETGA	459
Db	421	GLLFAVTSVAFVQMRQRHRRGT	KGVSVYRPAEVAETGA	459
RESULT 8				
US-10-723-795-2				
; Sequence 2, Application US/10723795				
; Publication No. US20040146955A1				
; GENERAL INFORMATION:				
; APPLICANT: Supuran, Claudiu				
; APPLICANT: Scozzafava, Andrea				
; APPLICANT: Pastorekova, Silvia				
; APPLICANT: Pastorek, Jaromir				
; TITLE OF INVENTION: CA IX-SPECIFIC INHIBITORS				
; FILE REFERENCE: MST-2393 US				
; CURRENT APPLICATION NUMBER: US/10/723,795				
; PRIOR FILING DATE: 2003-11-26				
; PRIOR APPLICATION NUMBER: 60/429,089				
; PRIOR FILING DATE: 2002-11-26				
; PRIOR APPLICATION NUMBER: 60/489,473				
; PRIOR FILING DATE: 2003-07-22				
; PRIOR APPLICATION NUMBER: 60/515,140				
; PRIOR FILING DATE: 2003-10-28				
; NUMBER OF SEQ ID NOS: 9				
; SOFTWARE: PatentIn version 3.2				
; SEQ ID NO 2				
; LENGTH: 459				
; TYPE: PRT				
; ORGANISM: Homo sapiens				
; US-10-723-795-2				

	Query Match	100.0%;	Score	2424;	DB	4;	Length	459;	
	Best Local Similarity	100.0%;	Pred.	No. 1.2e-173;					
	Matches	459;	Conservative	0;	Mismatches	0;	Indels	0; Gaps	0;
Qy	1	MAPICPSPWLPPLLIPAPAGGLTVOLLISLLLMPVHPORLPRMOEDSPLGGSGSGEDDPL	60						
Dd	1	MAPICPSPWLPPLLIPAPAGGLTVOLLISLLLMPVHPORLPRMOEDSPLGGSGSGEDDPL	60						
Qy	61	GEEIDLPSREDSPREDDPGSEDLPGCEEDLPGEEDLPVKPKSEBEGSLKLBDELPTVAAGP	120						
Dd	61	GEEIDLPSREDSPREDDPGSEDLPGCEEDLPGEEDLPVKPKSEBEGSLKLBDELPTVAAGP	120						
Qy	121	DPQPONNAHDKGDDQSHRYCGDPPWPVSVPACAGRFPQVDIRPOLAAFCPALRPL	180						
Dd	121	DPQPONNAHDKGDDQSHRYCGDPPWPVSVPACAGRFPQVDIRPOLAAFCPALRPL	180						
Qy	181	ELLGFQLPPELRRLNNGHSVOLTPPGLEMALPGGREYRALQLHLHWGAAGRPGSHT	240						
Dd	181	ELLGFQLPPELRRLNNGHSVOLTPPGLEMALPGGREYRALQLHLHWGAAGRPGSHT	240						
Qy	241	VEGHRFPAAIHHVHLSTAFARVDEALGRPGGLAVLAAPLEGPEENSAYEOLLSLEEIA	300						
Dd	241	VEGHRFPAAIHHVHLSTAFARVDEALGRPGGLAVLAAPLEGPEENSAYEOLLSLEEIA	300						
Qy	301	EESGSTQVPGLDISALLPSDFSRYFOYEGSLTTPPCAQGVIWTVFNQVMLSAKQLHTLS	360						
Dd	301	EESGSTQVPGLDISALLPSDFSRYFOYEGSLTTPPCAQGVIWTVFNQVMLSAKQLHTLS	360						
Qy	361	DTLWGPGDSRLQNFRATQPLNGRVIEASFPAGVDSSPRAEPVOLNSCLAAGDITIALVF	420						
Dd	361	DTLWGPGDSRLQNFRATQPLNGRVIEASFPAGVDSSPRAEPVOLNSCLAAGDITIALVF	420						

[illegible]

```

RESULT 10
US-10-723-860-461
; Sequence 461, Application US/10723860
; Publication No. US2004023606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagn
; TITLE OF INVENTION: Methods for Scr
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10723

```

APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
FILE REFERENCE: 05882.0193.NPUS01
CURRENT APPLICATION NUMBER: US/10/723,860

```
;
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 461
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-461

Query Match      100.0%; Score 2424; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e-173;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPILGGSGSGDDPL 60
DB 1 MAPLCSPWMLPLLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPILGGSGSGDDPL 60

QY 61 GEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

QY 121 DPOEPQNNARHDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
DB 121 DPOEPQNNARHDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180

QY 61 GEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

QY 121 DPOEPQNNARHDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
DB 121 DPOEPQNNARHDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180

QY 181 ELGFGQLPPLPELRLRNNHSHVQLTLPGLLEMALGPGREYRALQLHLHWGAAGRPGSEHT 240
DB 181 ELGFGQLPPLPELRLRNNHSHVQLTLPGLLEMALGPGREYRALQLHLHWGAAGRPGSEHT 240

QY 241 VEGHREPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
DB 241 VEGHREPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300

QY 301 EGESETQVPGLDISALLPSDFSRYFOYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
DB 301 EGESETQVPGLDISALLPSDFSRYFOYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360

QY 361 DTLWPGDSSLQLNFRATQPLNGRVIEASPPAGVDSPPRAAEVPQNSCLAAGDILALVF 420
DB 361 DTLWPGDSSLQLNFRATQPLNGRVIEASPPAGVDSPPRAAEVPQNSCLAAGDILALVF 420

QY 421 GLLFAVTSVAFVQMRQRHRRGKGGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFVQMRQRHRRGKGGVSYRPAEVAETGA 459

RESULT 11
US-10-795-933-6
; Sequence 6, Application US/10795933
; Publication No. US20040259126A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 465 California Street, Suite 450
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION NUMBER: US/10/795,933
; FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US/08/250,190
; PRIOR FILING DATE: 1994-06-15
; PRIOR APPLICATION NUMBER: 08/177,093
; PRIOR FILING DATE: 1993-12-30
; PRIOR APPLICATION NUMBER: 07/964,589
; PRIOR FILING DATE: 1992-10-21
; PRIOR APPLICATION NUMBER: PV-709-92
; PRIOR FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 459
; TYPE: PRT
; NAME: Lauder, Leona L.

;
; ORGANISM: HUMAN
; US-10-795-933-6

Query Match      100.0%; Score 2424; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e-173;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPLCSPWMLPLLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPILGGSGSGDDPL 60
DB 1 MAPLCSPWMLPLLIIPAPAGLTVQLLSLLLLMPVHPQRLPRMQEDSPILGGSGSGDDPL 60

QY 61 GEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

QY 121 DPOEPQNNARHDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
DB 121 DPOEPQNNARHDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180

QY 61 GEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
DB 61 GEDLPSEEDSPREEDPPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120

QY 121 DPOEPQNNARHDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
DB 121 DPOEPQNNARHDKEGDDQSHWRYGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180

QY 181 ELGFGQLPPLPELRLRNNHSHVQLTLPGLLEMALGPGREYRALQLHLHWGAAGRPGSEHT 240
DB 181 ELGFGQLPPLPELRLRNNHSHVQLTLPGLLEMALGPGREYRALQLHLHWGAAGRPGSEHT 240

QY 241 VEGHREPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300
DB 241 VEGHREPAEIHVVHLSTAFARVDEALGRPGGLAVLAFLAFLAFLAFLAFLAFLAFLA 300

QY 301 EGESETQVPGLDISALLPSDFSRYFOYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360
DB 301 EGESETQVPGLDISALLPSDFSRYFOYEGSLTTPPCAQGVITVFNQTVMLSAKQLHTLS 360

QY 361 DTLWPGDSSLQLNFRATQPLNGRVIEASPPAGVDSPPRAAEVPQNSCLAAGDILALVF 420
DB 361 DTLWPGDSSLQLNFRATQPLNGRVIEASPPAGVDSPPRAAEVPQNSCLAAGDILALVF 420

QY 421 GLLFAVTSVAFVQMRQRHRRGKGGVSYRPAEVAETGA 459
DB 421 GLLFAVTSVAFVQMRQRHRRGKGGVSYRPAEVAETGA 459

RESULT 12
US-10-888-694-2
; Sequence 2, Application US/10888694
; Publication No. US20050003425A1
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 465 California Street, Suite 450
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION NUMBER: US/10/888,694
; FILING DATE: 08-Jul-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/772,719
; FILING DATE: 30-Jan-2001
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
```


	Matches	459;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MAPLCPSFWLP	LLIPAPAPGLTVQLLSLLLMPVHPQR	LPMQEDSP	PLGGSSG	GEDDPL	60			
Db	1	MAPLCPSFWLP	LLIPAPAPGLTVQLLSLLLMPVHPQR	LPMQEDSP	PLGGSSG	GEDDPL	60			
Qy	61	GEEDLPSSEDS	PREDPPGEEDLPGCEEDLPEVKPKSBEESG	LKLEDLPTVEAPG	120					
Db	61	GEEDLPSSEDS	PREDPPGEEDLPGCEEDLPEVKPKSBEESG	LKLEDLPTVEAPG	120					
Qy	121	DPOEPQNNAUR	KDEGDQSHWRVYGGDPWPWPRVS	PACAGRFPQSDVIRPQLAAFCPALRPL	180					
Db	121	DPOEPQNNAUR	KDEGDQSHWRVYGGDPWPWPRVS	PACAGRFPQSDVIRPQLAAFCPALRPL	180					
Qy	181	ELIGFOLPPI	PELRNRNGHSVQLTPGLEMALPGREYVALOHLHWGAAGRP	SGSEHT	240					
Db	181	ELIGFOLPPI	PELRNRNGHSVQLTPGLEMALPGREYVALOHLHWGAAGRP	SGSEHT	240					
Qy	241	VEGHRFPFAE	IHVHLSLTAFAFVDALGRPGGLAVLAFLKEEGPEENSAYEQLLSRLEBIA	300						
Db	241	VEGHRFPFAE	IHVHLSLTAFAFVDALGRPGGLAVLAFLKEEGPEENSAYEQLLSRLEBIA	300						
Qy	301	BEGSETQVPG	LDTISALLPSDFSRFYQEGSLTTPPCAQGVITWFNQTWMLSAKQLHTLS	360						
Db	301	BEGSETQVPG	LDTISALLPSDFSRFYQEGSLTTPPCAQGVITWFNQTWMLSAKQLHTLS	360						
Qy	361	DTLWGPGDSL	RQLNFRATQPLNGRVIEASPPAGVDSSPRAAEPVOLNSCLAAGDILALVF	420						
Db	361	DTLWGPGDSL	RQLNFRATQPLNGRVIEASPPAGVDSSPRAAEPVOLNSCLAAGDILALVF	420						
Qy	421	GLLPVNTSVAF	LVMRRQHRGRTKGGSYRPAEVAETGA	459						
Db	421	GLLPVNTSVAF	LVMRRQHRGRTKGGSYRPAEVAETGA	459						

Search completed: February 12, 2006, 20:11:17
Job time : 179 secs

```

RESULT 15
US-10-794-514A-294
; Sequence 294, Application US/10794514A
; Publication No. US20050112134A1
; GENERAL INFORMATION:
; APPLICANT: Graddis, Thomas
; APPLICANT: Laus, Reiner
; APPLICANT: Diegel, Michael
; APPLICANT: Vidovic, Damir
; TITLE OF INVENTION: Compositions and Methods Employing Alternative
; TITLE OF INVENTION: Reading Frame Polypeptides for the Treatment of
; TITLE OF INVENTION: Cancer and Infectious Disease
; FILE REFERENCE: 11311.1003U
; CURRENT APPLICATION NUMBER: US/10/794,514A
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 733
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Derived From Human Sequence
US-10-794-514A-294

Query Match          100.0%; Score 2424; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e-173;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy      1  MAPLCPSPWLPLLIIPAPAPGLTVQLLLSILLMLPVHPQRLPRMQDSPLGGGSGSGDDPL 60
      |||
Db      1  MAPLCPSPWLPLLIIPAPAPGLTVQLLLSILLMLPVHPQRLPRMQDSPLGGGSGSGDDPL 60
      |||

Qy     61  GREDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
      |||
Db     61  GREDLPSEEDSPREDDPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLPGEEDLP 120
      |||

Qy    121  DPQEPONNAHRDKEGDGDSHWRYQGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
      |||
Db    121  DPQEPONNAHRDKEGDGDSHWRYQGGDPWPVRVSPACAGRFQSPVDIRPQLAAFCPALRPL 180
      |||

```


Result No.	Score	Query Match	Length	DB	ID	Description
1	2424	100.0	459	7	US-11-186-284-12	Sequence 12, Appl
2	562	23.2	337	6	US-10-063-703-74	Sequence 74, Appl
3	562	23.2	337	7	US-11-052-240-74	Sequence 74, Appl
4	439	18.1	255	7	US-11-054-281-304	Sequence 304, App
5	425.5	17.6	283	6	US-10-055-877-167	Sequence 167, App
6	402.5	16.6	260	6	US-10-055-877-163	Sequence 163, App
7	377	15.6	262	6	US-10-055-877-16	Sequence 16, Appl
8	377	15.6	262	6	US-10-055-877-162	Sequence 162, App
9	376	15.5	262	6	US-10-055-877-18	Sequence 18, Appl
10	362.5	15.0	261	6	US-10-055-877-165	Sequence 165, App
11	359.5	14.8	261	6	US-10-055-877-164	Sequence 164, App
12	355	14.6	421	6	US-10-673-781-1	Sequence 1, Appl
13	354	14.6	260	6	US-10-055-877-166	Sequence 166, App
14	325	13.4	1445	7	US-11-169-041-181	Sequence 181, Appl
15	304.5	12.6	374	7	US-11-097-728-4	Sequence 4, Appl
16	304.5	12.6	2314	7	US-11-097-728-2	Sequence 2, Appl
17	304.5	12.6	2353	7	US-11-097-728-6	Sequence 6, Appl
18	300	12.4	312	7	US-11-054-281-20	Sequence 20, Appl
19	300	12.4	312	7	US-11-054-281-86	Sequence 86, Appl
20	292.5	12.1	266	7	US-11-054-281-88	Sequence 88, Appl
21	290.5	12.0	294	7	US-11-054-281-87	Sequence 87, Appl
22	290	12.0	262	7	US-11-054-281-89	Sequence 89, Appl
23	277	11.4	328	6	US-10-131-826A-34	Sequence 34, Appl
24	263.5	10.9	328	6	US-10-131-826A-326	Sequence 326, App
25	245	10.1	252	6	US-10-467-657-276	Sequence 276, App

.....

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.


```

RESULT 5
US - 10-055-877-167
; Sequence 167, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: DeCristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerkusen, Bryan

```

Query Match	17.6%	Score	425.5	DB	6	Length	283
Best Local Similarity	36.1%	Pred. No.	1.3e-25				
Matches	105	Conservative	42	Mismatches	103	Indels	41
Gaps	11						
Qy	124	EPONNAHRDKEGDQSHWRYGDPMPRVSPA	-----CAGRFQSPDVRPQLAAFCPALR	178			
Db	10	EHSNNAH-----VLWHKLYPIANGNCQGERQSPINIQTKEAKYDPSLK	53				
Qy	179	PLELLGFLPPLPELRLRNNGHSVOLTLPGLLEMAL---	GP-----GREYRALQLHLHWGAA	232			
Db	54	PLS--LSYDAATAKEFEIVNNGHSFQVEFDDSDDKSVLSGGPLPAGHYRLKQFFHWGGA	112				
Qy	233	GR--PGSEHTEGHRFPAAIHHVHL--STAFARVDEALGRPGGLAVLAAPFEEGP--ENSA	288				
Db	113	SSDDQGSSEHTVDGKKYAAELHLVHWNSTKYSYKEAVSKPDGLAVGLVFLKVGQYQENPG	172				
Qy	289	YEQLLSLEIAEEGSETQVPGDLIDISALLPSDFSR--YFOYEGSLTTPPCAGQVITWTFNQ	347				
Db	173	LQKVWDALSSIKTKGKSATTFNPDFTLLPSEKLRDWTWYPGSLTTPPLETSVTWIVLKE	232				
Qy	348	TVMLSAQLHTLSDTLWG-----PG--DSRLQLNFRATQPLNGRVIEAST	390				

Query Match		15.0%; Score 362.5; DB 6; Length 261;
Best Local Similarity		32.6%; Pred. No. 7.5e-21;
Matches		87; Conservative 46; Mismatches 111; Indels 23; Gaps 8;
QY	132	DKEGDDQSHWRYGDDPPWPRVSPACAGRFOSPDVIRPQLAFCPALRPLELLGFLQPLP 191
DB	10	DKNQPEQ-----WSKLYPIANGNNQSPVDIKTSETKHTSLKPISV---SYNPAT 56
QY	192	ELRLNNGHVSQVLTLPPLGLEMAL---GP-GREYRALQLHLHWGAAGRPGSEHTVEGHRFP 247
DB	57	AKELIINVGHSHFVTFEDNDRSVLKGGLSDSYRLFOFHFWGSTNEHGSHTVDGVKYS 116
QY	248	AETHVHLSTA-FARVDEALGRGGLAVLAFTLEEGPEENSAYEQLLSRLEETAEBSSET 306
DB	117	AEHLHTHWSAKYSLSLAESAASAKADGLAVIGVLMKVG-EANPKLOKILDALQAIKTKGRA 175
QY	307	QVPGCLDISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQL---HTLSDTL 363
DB	176	PFTNFDSTLLPSSLD-FWTYPGSLTHPPLYESVTWICKESISVSSEQLAQFRSLLSNV 234
QY	364	WGPCDSRLQNLFRATQPLNGRVIEASF 390
DB	235	EGDNAVPMQHNRRPTQLKGRTRASF 261
RESULT 11		
US-10-055-877-164		
Sequence 164, Application US/10055877		
Publication No. US20050289241A1		
GENERAL INFORMATION:		
APPLICANT: Decristofaro, Marc		
APPLICANT: Padigaru, Muralidhara		
APPLICANT: Miller, Charles		
APPLICANT: Tchernev, Velizar		
APPLICANT: Zhong, Mei		
APPLICANT: Anderson, David		
APPLICANT: Ballinger, Robert		
APPLICANT: Gerlach, Valerie		
APPLICANT: Spytek, Kimberly		
APPLICANT: Ratelli, Luca		
APPLICANT: Kekuda, Ramesh		
APPLICANT: Guo, Xiaojia		
APPLICANT: Zerhuseen, Bryan		
APPLICANT: Andrew, David		
APPLICANT: Mezes, Peter		
APPLICANT: Patturajan, Meera		
APPLICANT: Burgess, Catherine		
APPLICANT: Eisen, Andrew		
APPLICANT: Wolenc, Adam		
APPLICANT: Baumgartner, Jason		
APPLICANT: Shinkets, Richard		
APPLICANT: Gusev, Vladimir		
APPLICANT: Vernet, Corine		
APPLICANT: Taupier Jr., Raymond		
APPLICANT: Pena, Carol		
APPLICANT: Shenoy, Suresh		
APPLICANT: Li, Li		
APPLICANT: Casman, Stacie		
APPLICANT: Boldog, Ference		
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby		
FILE REFERENCE: 21402-251		
CURRENT APPLICATION NUMBER: US/10/055,877		
CURRENT FILING DATE: 2002-01-22		
PRIOR APPLICATION NUMBER: 60/262,892		
PRIOR FILING DATE: 2001-01-19		
PRIOR APPLICATION NUMBER: 60/263,598		
PRIOR FILING DATE: 2001-01-23		
PRIOR APPLICATION NUMBER: 60/263,799		
PRIOR FILING DATE: 2001-01-24		
PRIOR APPLICATION NUMBER: 60/264,117		
PRIOR FILING DATE: 2001-01-25		
PRIOR APPLICATION NUMBER: 60/264,139		

Query Match		14.8%; Score 359.5; DB 6; Length 261;
Best Local Similarity		32.6%; Pred. No. 1.3e-20;
Matches		87; Conservative 45; Mismatches 112; Indels 23; Gaps 8;
QY	132	DKEGDDQSHWRYGDDPPWPRVSPACAGRFOSPDVIRPQLAFCPALRPLELLGFLQPLP 191
DB	10	DKNQPEQ-----WSKLYPIANGNNQSPVDIKTSETKHTSLKPISV---SYNPAT 56
QY	192	ELRLNNGHVSQVLTLPPLGLEMAL---GP-GREYRALQLHLHWGAAGRPGSEHTVEGHRFP 247
DB	57	AKELIINVGHSHFVTFEDNDRSVLKGGLSDSYRLFOFHFWGSTNEHGSHTVDGVKYS 116
QY	248	AETHVHLSTA-FARVDEALGRGGLAVLAFTLEEGPEENSAYEQLLSRLEETAEBSSET 306
DB	117	AEHLHTHWSAKYSLSLAESAASAKADGLAVIGVLMKVG-EANPKLOKILDALQAIKTKGRA 175
QY	307	QVPGCLDISALLPSDFSRYFOYEGSLTTPPCAQQGVITVFNQTVMLSAKQL---HTLSDTL 363
DB	176	PFTNFDSTLLPSSLD-FWTYPGSLTHPPLYESVTWICKESISVSSEQLAQFRSLLSNV 234
QY	364	WGPCDSRLQNLFRATQPLNGRVIEASF 390
DB	235	EGDNAVPMQHNRRPTQLKGRTRASF 261
RESULT 12		
US-10-673-781-1		
Sequence 1, Application US/10673781		
Publication No. US2005026089A1		
GENERAL INFORMATION:		
APPLICANT: Qirwei Shi		
TITLE OF INVENTION: Differential Immunoassay		
FILE REFERENCE: 1112-1-080N		
CURRENT APPLICATION NUMBER: US/10/673,781		
CURRENT FILING DATE: 2003-09-29		
PRIOR APPLICATION NUMBER: US/09/938,270B		
PRIOR FILING DATE: 2001-08-23		
PRIOR APPLICATION NUMBER: US 60/227,536		
PRIOR FILING DATE: 2000-08-24		
PRIOR APPLICATION NUMBER: US 60/292,497		
PRIOR FILING DATE: 2001-05-21		
NUMBER OF SEQ ID NOS: 2		
SOFTWARE: FastSeq for Windows Version 4.0		
SEQ ID NO 1		
LENGTH: 421		
TYPE: PRT		
ORGANISM: homosapien		
US-10-673-781-1		
Query Match		14.6%; Score 355; DB 6; Length 421;
Best Local Similarity		30.1%; Pred. No. 5.2e-20;
Matches		94; Conservative 55; Mismatches 137; Indels 26; Gaps 9;

THIS PAGE BLANK (USPTO)